J. Differential Equations 252 (2012) 1–34

Contents lists available at SciVerse ScienceDirect

Journal of Differential Equations

www.elsevier.com/locate/jde

# Invariant tori of full dimension for a nonlinear Schrödinger equation ☆

Jiansheng Geng

*Department of Mathematics, Nanjing University, Nanjing 210093, PR China*

ARTICLE INFO

*Article history:*
Received 2 September 2009
Revised 3 September 2011
Available online 28 September 2011

ABSTRACT

In this paper, we consider the one-dimensional nonlinear Schrödinger equation

$$\mathrm{i}u_t - u_{xx} + mu + f\left(|u|^2\right)u = 0$$

with periodic boundary conditions or Dirichlet boundary conditions, where $f$ is a real analytic function in some neighborhood of the origin satisfying $f(0) = 0$, $f'(0) \neq 0$. We prove that for each given constant potential $m$, when the frequencies, as a function of the amplitudes, can be regarded as the independent parameters, the equation admits a Whitney smooth family of small-amplitude, time almost-periodic solutions with all frequencies. The proof is based on a Birkhoff normal form reduction and an improved version of the KAM theorem.



## 1. Introduction and main result

At present, there have been many remarkable results in KAM (Kolmogorov–Arnold–Moser) theory of Hamiltonian partial differential equations (see [1,9,11–22,25–27,30–33] for KAM methods, and see [2–6,10] for CWB (Craig–Wayne–Bourgain) methods). In the above papers, the authors proved the existence of quasi-periodic solutions, which is to say, the persistence of finite dimensional invariant tori, under Hamiltonian perturbation. With respect to the case of almost-periodic solutions, meaning the persistence of infinite dimensional invariant tori under Hamiltonian perturbation, up to now there

☆ The work is partially supported by NSFC grant 10771098 and 973 projects of China 2007CB814800. This work is also partially supported by Program for New Century Excellent Talents in University.

*E-mail address:* jgeng@nju.edu.cn.

0022-0396/$ – see front matter 
doi:10.1016/j.jde.2011.09.006

are only a few results available. Precisely speaking, in [7], Bourgain considered the nonlinear wave equation of the form

$$u_{tt} - u_{xx} + V(x)u + \varepsilon F(u) = 0,$$

under Dirichlet boundary conditions

$$u(t,0) = u(t,\pi) = 0,$$

and proved that, for "typical" periodic potentials $V(x)$, the above equations admit invariant tori of full dimension in the neighborhood of $u = 0$. In [28], Pöschel considered nonlinear Schrödinger equations of the form

$$\mathrm{i}u_t - u_{xx} + V(x)u + \Psi\big(f\big(|\Psi u|^2\big)\Psi u\big) = 0,$$
$$u(t,0) = u(t,\pi) = 0,$$
$$\Psi : H_0^s\big([0,\pi]\big) \to H_0^{s+\sigma}\big([0,\pi]\big), \quad \sigma > \frac{1}{4},$$

and proved that, for "almost all" potentials $V \in L^2([0,\pi])$, the above equations admit uncountably many small-amplitude almost-periodic solutions (see also [24] for the case of the higher dimensional beam equations). In particular, at the very end of his paper, Pöschel remarked that, "the problem is greatly simplified by the assumption that some potential is available serving as an infinite dimensional parameter. This decouples the problem of choosing amplitudes for the action coordinates and of adjusting the frequencies. Nothing is known, however, about the existence of almost-periodic solutions for a nonlinear Schrödinger equation such as $\mathrm{i}u_t - u_{xx} + mu + f(|u|^2)u = 0$ with Dirichlet boundary conditions on $[0,\pi]$, although, for example, a complete non-degenerate Birkhoff normal form up to order four is available [22]."

In [8], Bourgain considered the nonlinear Schrödinger equations with periodic boundary conditions

$$\mathrm{i}u_t - u_{xx} + Mu + f\big(|u|^2\big)u = 0,$$

where $M$ is a random Fourier multiplier defined by

$$\widehat{Mu}(n) = V_n \hat{u}(n)$$

and $(V_n)_{n\in\mathbb{Z}}$ are independently chosen in $[-1,1]$, and proved that, for appropriate $M$, the above equation has an invariant torus of full dimension with the solution satisfying slower decay than those of [7,28,24]. Moreover, Bourgain also pointed out that, "the multiplier $M = (V_n)$ is to be considered as a parameter and the role of this parameter is essential to ensure appropriate nonresonance properties of the modulated frequencies along the iteration. In the absence of exterior parameters, these conditions need to be realized from amplitude-frequency modulation and suitable restriction of the action-variables. This problem is harder. Indeed, a fast decay of the action-variables (enhancing convergence of the process) allows less frequency modulation and worse small divisors."

In this paper, we try to address the open problem in [28,8]. We will consider the nonlinear Schrödinger equation

$$\mathrm{i}u_t - u_{xx} + mu + f\big(|u|^2\big)u = 0, \tag{1.1}$$

under periodic boundary conditions

$$u(t,x) = u(t,x+2\pi), \tag{1.2}$$

or Dirichlet boundary conditions

$$u(t,0)=0=u(t,\pi), \tag{1.3}$$

where $f$ is a real analytic function in some neighborhood of the origin with $f(0)=0$, $f'(0)\neq 0$. As mentioned before, we will extract the parameters from amplitude-frequency modulation. One of the main difficulties is that if all the modulated frequencies are treated at the same time, because there will be infinitely many tangential frequencies, we do not know how to establish measure estimates for retained frequencies. Thus in our approach to this problem, at every KAM step we treat finitely many tangential frequencies, with the number of tangential frequencies under control. After this KAM step, our difficulty is to continue to extract parameters from the remaining normal coordinates. To do this, further action-angle transformations are needed, where generally speaking, the new perturbation is at most of order three in the normal coordinates. At the same time, the new modulated frequencies need to be realized from the terms of order four in the normal coordinates, which is impossible. To overcome this difficulty, we modify the previous KAM mechanism and augment the small divisor conditions so that our new perturbation is about of order five in the normal coordinates, whereupon the next action-angle transformation is feasible. To obtain the above normal form, we will deal with the small divisors with three or four normal frequencies, thus our small divisor conditions are significantly different from the previous ones, and hence our measure estimates are more complicated. To fulfill this more complicated measure estimates, we consider the perturbations having a compact form (periodic boundary conditions) or perturbations having decay property (Dirichlet boundary conditions) and take the normal frequencies increasing super-linearly (see Section 5.3). If the perturbations are of order five and the normal form has the fourth order average terms, as in the first KAM step, the induction step involving action-angle transformations is applicable.

Our main results can be stated as follows.

**Theorem 1.** *Consider the one-dimensional nonlinear Schrödinger equation*

$$\mathrm{i}u_t-u_{xx}+mu+f\big(|u|^2\big)u=0,$$

*with periodic boundary conditions*

$$u(t,x+2\pi)=u(t,x),$$

*where $f$ is a real analytic function in some neighborhood of the origin with $f(0)=0$, $f'(0)\neq 0$. The linearized equation has solutions*

$$u(t,x)=\sum_{n\in\mathbb{Z}}\sqrt{\xi_n}e^{\mathrm{i}((n^2+m)t+nx)},\quad 0<\xi_n\ll 1,$$

*taking $\xi=(\dots,\xi_n,\dots)_{n\in\mathbb{Z}}\in\mathcal{O}$ as parameters, when the frequencies, as a function of $\xi$, can be regarded as the independent parameters, there exists a positive-measure Cantor subset $\tilde{\mathcal{O}}\subset\mathcal{O}$, such that for any $\xi\in\tilde{\mathcal{O}}$, the above nonlinear equation has a solution*

$$u(t,x)=\sum_{n\in\mathbb{Z}}\sqrt{\xi_n}e^{\mathrm{i}(\omega_n t+nx)}+O\big(|\xi|^{\frac{3}{2}}\big),$$

$$\omega_n=n^2+m+O\big(|\xi|\big),\quad n\in\mathbb{Z}.$$

**Theorem 2.** *Consider the one-dimensional nonlinear Schrödinger equation*

$$\mathrm{i}u_t-u_{xx}+mu+f\big(|u|^2\big)u=0,$$

*with Dirichlet boundary conditions*

$$u(t,0)=0=u(t,\pi),$$

*where $f$ is a real analytic function in some neighborhood of the origin with $f(0)=0$, $f'(0)\neq 0$. The linearized equation has solutions*

$$u(t,x)=\sum_{n>0}\sqrt{\xi_n}e^{\mathrm{i}(n^2+m)t}\sin(nx),\quad 0<\xi_n\ll 1,$$

*taking $\xi=(\dots,\xi_n,\dots)_{n\in\mathbb{Z}_+}\in\mathcal{O}$ as parameters, when the frequencies, as a function of $\xi$, can be regarded as the independent parameters, there exists a positive-measure Cantor subset $\tilde{\mathcal{O}}\subset\mathcal{O}$, such that for any $\xi\in\tilde{\mathcal{O}}$, the above nonlinear equation has a solution*

$$u(t,x)=\sum_{n>0}\sqrt{\xi_n}e^{\mathrm{i}\omega_n t}\sin(nx)+O\big(|\xi|^{\frac{3}{2}}\big),$$
$$\omega_n=n^2+m+O\big(|\xi|\big),\quad n>0.$$

**Remark 1.** Under Dirichlet boundary conditions, the nonlinearity may be allowed to depend explicitly on the space variable $x$ in the real analytic way, but the proof is more complicated, we do not pursue this point. In this paper, to focus on the main ideas, we prove Theorem 1 in detail. We only clarify the differences between the proof of Theorem 2 and the proof of Theorem 1 at some appropriate places. Under periodic boundary conditions, if the nonlinearity depends explicitly on the space variable $x$, then there will be non-integrable terms in the normal form, I think, which should be an obstacle for proving the existence of full dimensional invariant tori.

**Remark 2.** Under Dirichlet boundary conditions, when the nonlinearity $f(|u|^2,x)=cg(x)|u|^2+\text{h.o.t.}$ ($c\neq 0$), from the following proof, it is easy to see that we can also get the existence of quasi-periodic solutions for the above nonlinear Schrödinger equation. In fact, by appropriately selecting the tangential site $\{i,j\}$, then $i^2+j^2-k^2-l^2\neq 0$ and $i^2-j^2+k^2-l^2\neq 0$. The interested readers can refer to [23,16,17] for similar techniques.

**Remark 3.** The measure of $\tilde{\mathcal{O}}\subset\mathcal{O}$ is positive, means that, for a fixed $\varepsilon>0$ small enough,

$$\mathcal{O}=\big\{\xi=(\dots,\xi_n,\dots)_{n\in\mathbb{Z}}\colon\ \xi_n\in\mathcal{O}_n=\big[\varepsilon e^{-2^{|n|}},2\varepsilon e^{-2^{|n|}}\big]\big\},$$

then

$$\frac{\operatorname{meas}(\mathcal{O}\setminus\tilde{\mathcal{O}})}{\operatorname{meas}(\mathcal{O})}=\max_{n\in\mathbb{Z}}\frac{\operatorname{meas}(\mathcal{O}_n\setminus\tilde{\mathcal{O}}_n)}{\operatorname{meas}(\mathcal{O}_n)}\to 0,\quad \text{as } \varepsilon\to 0.$$

**Remark 4.** From Remark 3 above, one can observe that the actions decay super-exponentially, it is certainly interesting whether the actions can be proved to decay exponentially like those of [8]. In addition, from the statement of the theorems, we add an assumption "the frequencies, as a function of $\xi$, can be regarded as the independent parameters", so that our measure estimates are feasible. Without the above assumption, in this paper, we do not know how to prove the above theorems.

The rest of the paper is devoted to the proof of Theorem 1. Section 2 is a preliminary section in which we define the weighted norms and compact forms and study their basic properties. In Section 3, we derive an integrable Birkhoff normal form of order four for the lattice Hamiltonian associated with (1.1) and (1.2), and then transform it into a parameterized Hamiltonian normal form. In

Section 4, we give details for one step of the KAM iteration. The proof of the theorem is completed in Section 5 by showing an iteration lemma, giving a convergence result, and finally conducting measure estimates.

## 2. Preliminary

### 2.1. Weighted norms

For a given $\rho > 0$, we let $\ell^\rho$ be the Banach space of bi-infinite, complex-valued sequences $q = (\{q_n\})$, endowed with the weighted norm

$$\|q\|_\rho = \sum_{n\in\mathbb{Z}} |q_n| e^{|n|\rho}.$$

Similarly, let $(\{\phi_n(x)\})$ be complete orthonormal basis in $L^2(T)$ and let $\mathcal{L}^\rho$ be the Banach space of functions $u(x) = \sum_{n\in\mathbb{Z}} q_n\phi_n(x)$ for $(\{q_n\}) \in \ell^\rho$, endowed with the norm $\|u\|_\rho = \|q\|_\rho$. Then $\mathcal{L}^\rho$ and $\ell^\rho$ are isometric, and the product of two functions $u(x) = \sum_{n\in\mathbb{Z}} p_n\phi_n(x)$, $v(x) = \sum_{n\in\mathbb{Z}} q_n\phi_n(x)$ in $\mathcal{L}^\rho$ defines the convolution $q * p$: $(q * p)_n = \sum_m q_{n-m} p_m$, $n \in \mathbb{Z}$, in $\ell^\rho$, under which $\ell^\rho$ becomes a Banach algebra. In particular,

$$\|q * p\|_\rho \leqslant \|q\|_\rho \|p\|_\rho,$$

for any $p, q \in \ell^\rho$.

Let $|\cdot|$ denote the sup-norm of complex vectors. For given $r, s > 0$, we let $D(r, s)$ be the complex neighborhood

$$D(r,s) = \bigl\{(\theta, I, q)\colon\ |\mathrm{Im}\,\theta| = \bigl|\mathrm{Im}(\theta_1, \dots, \theta_b)\bigr| < r,\ |I| = \bigl|(I_1, \dots, I_b)\bigr| < s^2,\ \|q\|_\rho < s\bigr\}$$

of $\mathbb{T}^b \times \{I = 0\} \times \{q = 0\}$ in $\mathbb{T}^b \times \mathbb{R}^b \times \ell^\rho$. Also let $\mathcal{O}$ be a bounded set in $\mathbb{R}^b_+$.

Let $F(\theta, I, q, \bar{q})$ be a real analytic function on $D(r, s)$ which depends on a parameter $\xi \in \mathcal{O}$ Whitney smoothly (i.e., $C^1$ in the sense of Whitney). We expand $F$ into the Taylor–Fourier series with respect to $\theta$, $I$, $q$, $\bar{q}$:

$$F(\theta, I, q, \bar{q}) = \sum_{\alpha,\beta} F_{\alpha\beta} q^\alpha \bar{q}^\beta,$$

where, for multi-indices $\alpha \equiv (\dots, \alpha_n, \dots)$, $\beta \equiv (\dots, \beta_n, \dots)$, $\alpha_n, \beta_n \in \mathbb{N}$ with finitely many non-vanishing components,

$$F_{\alpha\beta} = \sum_{k\in\mathbb{Z}^b,\, l\in\mathbb{N}^b} F_{kl\alpha\beta}(\xi) I^l e^{\mathrm{i}\langle k,\theta\rangle}.$$

We define the weighted norm of $F$ as

$$\|F\|_{D(r,s),\mathcal{O}} \equiv \sup_{\substack{\|q\|_\rho < s\\ \|\bar{q}\|_\rho < s}} \sum_{\alpha,\beta} \|F_{\alpha\beta}\| \bigl|q^\alpha\bigr| \bigl|\bar{q}^\beta\bigr|,$$

where

$$\|F_{\alpha\beta}\| \equiv \sum_{k,l} |F_{kl\alpha\beta}|_{\mathcal{O}} s^{2|l|} e^{|k|r}, \qquad |F_{kl\alpha\beta}|_{\mathcal{O}} \equiv \sup_{\xi\in\mathcal{O}} \biggl(|F_{kl\alpha\beta}| + \biggl|\frac{\partial F_{kl\alpha\beta}}{\partial \xi}\biggr|\biggr).$$

In the above and for the rest of the paper, derivatives with respect to the parameter $\xi \in \mathcal{O}$ are understood in the sense of Whitney.

For a vector-valued function $G : D(r,s) \times \mathcal{O} \to \mathbb{C}^m$, $m < \infty$, we define its weighed norm by

$$\|G\|_{D(r,s),\mathcal{O}} \equiv \sum_{i=1}^{m} \|G_i\|_{D(r,s),\mathcal{O}}.$$

For the Hamiltonian vector field

$$X_F = \big(F_I, -F_\theta, \{\mathrm{i}F_{q_n}\}, \{-\mathrm{i}F_{\bar{q}_n}\}\big)$$

associated with a function $F$ on $D(r,s) \times \mathcal{O}$, we define its weighted norm by

$$\|X_F\|_{D(r,s),\mathcal{O}} \equiv \|F_I\|_{D(r,s),\mathcal{O}} + \frac{1}{s^2}\|F_\theta\|_{D(r,s),\mathcal{O}} + \frac{1}{s}\bigg(\sum_n \|F_{q_n}\|_{D(r,s),\mathcal{O}} e^{|n|\rho} + \sum_n \|F_{\bar{q}_n}\|_{D(r,s),\mathcal{O}} e^{|n|\rho}\bigg).$$

Let $F, G$ be two real analytic functions on $D(r,s)$ which depend on a parameter $\xi \in \mathcal{O}$ Whitney smoothly.

**Lemma 2.1.**

$$\|FG\|_{D(r,s),\mathcal{O}} \leqslant \|F\|_{D(r,s),\mathcal{O}} \|G\|_{D(r,s),\mathcal{O}}.$$

**Proof.** Since $(FG)_{kl\alpha\beta} = \sum_{k',l',\alpha',\beta'} F_{k-k',l-l',\alpha-\alpha',\beta-\beta'} G_{k'l'\alpha'\beta'}$, we have

$$\begin{aligned}
\|FG\|_{D(r,s),\mathcal{O}} &= \sup_{\substack{\|q\|_\rho < s \\ \|\bar{q}\|_\rho < s}} \sum_{k,l,\alpha,\beta} \big|(FG)_{kl\alpha\beta}\big|_{\mathcal{O}} s^{2l} \big|q^\alpha\big| \big|\bar{q}^\beta\big| e^{|k|r} \\
&\leqslant \sup_{\substack{\|q\|_\rho < s \\ \|\bar{q}\|_\rho < s}} \sum_{k,l,\alpha,\beta} \sum_{k',l',\alpha',\beta'} |F_{k-k',l-l',\alpha-\alpha',\beta-\beta'} G_{k'l'\alpha'\beta'}|_{\mathcal{O}} s^{2l} \big|q^\alpha\big| \big|\bar{q}^\beta\big| e^{|k|r} \\
&\leqslant \|F\|_{D(r,s),\mathcal{O}} \|G\|_{D(r,s),\mathcal{O}}. \qquad \square
\end{aligned}$$

**Lemma 2.2** *(Cauchy inequalities).*

$$\|F_\theta\|_{D(r-\sigma,s),\mathcal{O}} \leqslant \frac{1}{\sigma}\|F\|_{D(r,s),\mathcal{O}},$$

$$\|F_I\|_{D(r,\frac{1}{2}s),\mathcal{O}} \leqslant \frac{4}{s^2}\|F\|_{D(r,s),\mathcal{O}},$$

*and*

$$\|F_{q_n}\|_{D(r,\frac{1}{2}s),\mathcal{O}} \leqslant \frac{2}{s}\|F\|_{D(r,s),\mathcal{O}} e^{|n|\rho},$$

$$\|F_{\bar{q}_n}\|_{D(r,\frac{1}{2}s),\mathcal{O}} \leqslant \frac{2}{s}\|F\|_{D(r,s),\mathcal{O}} e^{|n|\rho}.$$

**Proof.** It follows from the standard Cauchy estimate. $\square$

### 2.2. Compact form

Consider the Poisson bracket

$$\{F,G\}=\left\langle\frac{\partial F}{\partial I},\frac{\partial G}{\partial\theta}\right\rangle-\left\langle\frac{\partial F}{\partial\theta},\frac{\partial G}{\partial I}\right\rangle+\mathrm{i}\sum_n\left(\frac{\partial F}{\partial q_n}\frac{\partial G}{\partial\bar q_n}-\frac{\partial F}{\partial\bar q_n}\frac{\partial G}{\partial q_n}\right).$$

A real analytic function

$$F=F(\theta,I,q,\bar q)=\sum_{k,\alpha,\beta}F_{k\alpha\beta}(I)e^{\mathrm{i}\langle k,\theta\rangle}q^\alpha\bar q^\beta$$

on $D(r,s)$ is said to admit a *compact form* if

$$F_{k\alpha\beta}=0,\quad\text{whenever}\ \sum_{j=1}^b k_j j+\sum_n(\alpha_n-\beta_n)n\neq 0,$$

where $k=(k_1,k_2,\dots,k_b)\in\mathbb{Z}^b$, and $\alpha\equiv(\dots,\alpha_n,\dots)$, $\beta\equiv(\dots,\beta_n,\dots)$, $\alpha_n,\beta_n\in\mathbb{N}$, with finitely many non-vanishing components.

A crucial idea in the proof of our main result is to show that compact forms will be preserved by KAM iterations. This property, enabling the consideration of essentially finite small divisors in each KAM step, will play an important role in the measure estimates later on.

**Lemma 2.3.** *Consider two real analytic functions $F(\theta,I,q,\bar q)$, $G(\theta,I,q,\bar q)$ on $D(r,s)$. If both $F$ and $G$ have compact forms, then so does $\{F,G\}$.*

**Proof.** Let

$$F=\sum_{A_1}F_{k_1\alpha_1\beta_1}(I)e^{\mathrm{i}\langle k_1,\theta\rangle}q^{\alpha_1}\bar q^{\beta_1},$$
$$G=\sum_{A_2}G_{k_2\alpha_2\beta_2}(I)e^{\mathrm{i}\langle k_2,\theta\rangle}q^{\alpha_2}\bar q^{\beta_2},$$

where

$$A_i=\left\{(k_i,\alpha_i,\beta_i)\colon\ \sum_{j=1}^b k_{ij}j+\sum_n(\alpha_{in}-\beta_{in})n=0\right\},$$
$$k_i=(k_{i1},k_{i2},\dots,k_{ib}),$$
$$\alpha_i=\big(\{\alpha_{in}\}_{n\in\mathbb{Z}}\big),$$
$$\beta_i=\big(\{\beta_{in}\}_{n\in\mathbb{Z}}\big),$$

for $i=1,2$ respectively. A straightforward calculation yields that

$$\{F,G\}=\sum_{A_1}\sum_{A_2}\left\langle\frac{\partial F_{k_1\alpha_1\beta_1}(I)}{\partial I},\mathrm{i}k_2\right\rangle G_{k_2\alpha_2\beta_2}(I)e^{\mathrm{i}\langle k_1,\theta\rangle}q^{\alpha_1}\bar q^{\beta_1}e^{\mathrm{i}\langle k_2,\theta\rangle}q^{\alpha_2}\bar q^{\beta_2}$$

$$-\sum_{A_1}\sum_{A_2}\left\langle \mathrm{i}k_1, \frac{\partial G_{k_2\alpha_2\beta_2}(I)}{\partial I}\right\rangle F_{k_1\alpha_1\beta_1}(I)e^{\mathrm{i}\langle k_1,\theta\rangle}q^{\alpha_1}\bar{q}^{\beta_1}e^{\mathrm{i}\langle k_2,\theta\rangle}q^{\alpha_2}\bar{q}^{\beta_2}$$
$$+\,\mathrm{i}\sum_m \sum_{\tilde{A}_1\cup\tilde{A}_2} F_{k_1\alpha_1\beta_1}(I)G_{k_2\alpha_2\beta_2}(I)e^{\mathrm{i}\langle k_1,\theta\rangle}e^{\mathrm{i}\langle k_2,\theta\rangle}q^{\alpha_1-e_m}\bar{q}^{\beta_1}q^{\alpha_2}\bar{q}^{\beta_2-e_m}$$
$$-\,\mathrm{i}\sum_m \sum_{\tilde{A}_1\cup\tilde{A}_2} F_{k_1\alpha_1\beta_1}(I)G_{k_2\alpha_2\beta_2}(I)e^{\mathrm{i}\langle k_1,\theta\rangle}e^{\mathrm{i}\langle k_2,\theta\rangle}q^{\alpha_1}\bar{q}^{\beta_1-e_m}q^{\alpha_2-e_m}\bar{q}^{\beta_2},$$

where for each $i=1,2$, $m\in\mathbb{Z}$, $e_m$ is the multi-index whose $m$th component is 1 and other components are all 0,

$$\tilde{A}_i=\tilde{A}_i(m)=\left\{(k_i,\alpha_i,\beta_i)\colon \sum_{j=1}^{b}k_{ij}j+(\alpha_{im}-\beta_{im})m+\sum_{n\in\mathbb{Z}\setminus\{m\}}(\alpha_{in}-\beta_{in})n=0\right\},$$
$$k_i=(k_{i1},k_{i2},\ldots,k_{ib}),$$
$$\alpha_i=\left(\{\alpha_{in}\}_{n\in\mathbb{Z}\setminus\{m\}}\right),$$
$$\beta_i=\left(\{\beta_{in}\}_{n\in\mathbb{Z}\setminus\{m\}}\right).$$

Since all terms above have compact forms, so does $\{F,G\}$. $\square$

## 3. Normal form

Using the Hamiltonian formulation, we re-write Eq. (1.1) with the periodic boundary conditions (1.2) as the Hamiltonian system

$$u_t=\mathrm{i}\frac{\partial H}{\partial \bar{u}},$$

where

$$H=\int_0^{2\pi}\left(|u_x|^2+m|u|^2\right)dx+\int_0^{2\pi}g\left(|u|^2\right)dx,$$

where $g$ is a primitive function of $f$.

Note that the operator $A=-\partial_{xx}+m$ with the periodic boundary condition has an orthonormal basis $\{\phi_n(x)=\sqrt{\frac{1}{2\pi}}e^{\mathrm{i}nx}\}$ and corresponding eigenvalues

$$\mu_n=n^2+m.$$

Let

$$u(x,t)=\sum_{n\in\mathbb{Z}}q_n(t)\phi_n(x).$$

Then associated with the symplectic structure $\mathrm{i}\sum_n dq_n\wedge d\bar{q}_n$, $\{q_n\}_{n\in\mathbb{Z}}$ satisfies the Hamiltonian equations

$$\dot{q}_n=\mathrm{i}\frac{\partial H}{\partial \bar{q}_n},\quad n\in\mathbb{Z},$$

where

$$H = \Lambda + G \tag{3.1}$$

with

$$\Lambda = \sum_{n\in\mathbb{Z}} \mu_n |q_n|^2,$$

$$G = \int_0^{2\pi} g\left(\left|\sum_{n\in\mathbb{Z}} q_n\phi_n\right|^2\right) dx.$$

**Lemma 3.1.** *The gradient $G_{\bar{q}}$ is a real analytic map from a neighborhood of the origin of $\ell^\rho$ into $\ell^\rho$, with*

$$\|G_{\bar{q}}\|_\rho = O\left(\|q\|_\rho^3\right).$$

**Proof.** Let $G_{\bar{q}} = (\{\frac{\partial G}{\partial \bar{q}_n}\})$, where

$$\frac{\partial G}{\partial \bar{q}_n} = \int_0^{2\pi} f\left(|u|^2\right)u\bar{\phi}_n\, dx$$

for $u = \sum_{n\in\mathbb{Z}} q_n\phi_n$, i.e., $\frac{\partial G}{\partial \bar{q}_n} = (f(|u|^2)u)_n$. Hence,

$$\|G_{\bar{q}}\|_\rho = \left\| f\left(|u|^2\right)u\right\|_\rho \leqslant c\|u\|_\rho^3 = c\|q\|_\rho^3.$$

The analyticity of $G_{\bar{q}}$ follows from the regularity of its components and its local boundedness [29, Appendix A]. □

Since $f(|u|^2)$ is real analytic in $|u|^2$, then $g(|u|^2)$ is real analytic in $|u|^2$. Making use of $u(x,t) = \sum_{n\in\mathbb{Z}} q_n(t)\phi_n(x)$ again, we may re-write $g$ as follows:

$$g\left(|u|^2\right) = \sum_{\alpha,\beta} g_{\alpha\beta} q^\alpha \bar{q}^\beta \phi^\alpha \bar{\phi}^\beta,$$

hence

$$G(q,\bar{q}) \equiv \int_0^{2\pi} g\left(\left|\sum_{n\in\mathbb{Z}} q_n(t)\phi_n(x)\right|^2\right) dx = \sum_{\alpha,\beta} G_{\alpha\beta} q^\alpha \bar{q}^\beta,$$

$$G_{\alpha\beta} = 0, \quad \text{if} \sum_{n\in\mathbb{Z}^d} (\alpha_n - \beta_n)n \neq 0. \tag{3.2}$$

Next, we make use of Taylor's expansion $f(|u|^2) = c|u|^2 + \text{h.o.t.}$, without loss of generality, we assume $c = 1$ for simplicity, then

$$G = \frac{1}{2}\sum_{i,j,k,l} G_{ijkl} q_i q_j \bar{q}_k \bar{q}_l + \text{h.o.t.},$$

where

$$G_{ijkl}=\int_0^{2\pi}\phi_i\phi_j\bar{\phi}_k\bar{\phi}_l\,dx.$$

In terms of [22,15], we have the following proposition.

**Proposition 1.** *For the Hamiltonian $H=\Lambda+G$, there exists a real analytic, symplectic change of coordinates $\Gamma$ in a neighborhood of the origin of $\ell^\rho$ which transforms the Hamiltonian $H$ into its Birkhoff normal form up to order four, specifically*

$$H\circ\Gamma=\Lambda+\bar{G}+K \tag{3.3}$$

*such that the corresponding Hamiltonian vector fields $X_{\bar{G}}$ and $X_K$ are real analytic in a neighborhood of the origin in $\ell^\rho$, where*

$$\bar{G}=\frac{1}{2}\sum_{i,j}\bar{G}_{ij}|q_i|^2|q_j|^2,\qquad \bar{G}_{ij}=\frac{2-\delta_{ij}}{2\pi},$$

$$|K|=O\big(\|q\|_\rho^6\big).$$

*Moreover, $K(q,\bar{q})$ has a compact form.*

For the proof see [22,15].

It is elementary to observe that $\int_0^{2\pi}|u|^2\,dx$ is an integral of nonlinear Schrödinger equation (1.1) as well as of its linear and nonlinear part separately. Accordingly, $\sum_n|q_n|^2$ is an integral of $H=\Lambda+G$, which also commutes with $\Lambda$ and $G$ individually. The normalizing transformation $\Gamma$ of Proposition 1 preserves this property, i.e., $\sum_n|q_n|^2\circ\Gamma=\sum_n|q_n|^2$. Hence $\sum_n|q_n|^2=C$ is another conserved quantity, then

$$\bar{G}=\frac{1}{2\pi}\bigg[\bigg(\sum_n|q_n|^2\bigg)^2-\frac{1}{2}\sum_n|q_n|^4\bigg]=\frac{1}{2\pi}C^2-\frac{1}{4\pi}\sum_n|q_n|^4.$$

Moreover, substitute $u=e^{\mathrm{i}\sigma t}v$ into nonlinear Schrödinger equation (1.1), then (1.1) will become

$$\mathrm{i}v_t-v_{xx}+(m-\sigma)v+f\big(|v|^2\big)v=0,$$

hence without loss of generality, we may fix $m=\frac{1}{5}$ in the following context. Thus the Hamiltonian (3.3) becomes (the inessential constants are suppressed)

$$H\circ\Gamma=\Lambda-\frac{1}{4\pi}\sum_n|q_n|^4+K. \tag{3.4}$$

Next, fix $|n|=0$ (in fact, under periodic boundary conditions, 0 can be replaced by a fixed positive integer $N$, but at $\nu$th step, the number of the tangential frequencies should have the upper bound $\nu N$), we introduce action-angle variables and parameters to the Birkhoff normal form (3.4). Let $\xi=(\dots,\xi_n,\dots)_{|n|=0}$ (the reason of writing this form is in accord with the following notations) be a parameter and $(I,\theta)$ be the standard action-angle variables in the $(\dots,q_n,\bar q_n,\dots)_{|n|=0}$-space around $\xi$. Then

$$q_n=\sqrt{I_n+\xi_n}e^{\mathrm{i}\theta_n},\qquad \bar q_n=\sqrt{I_n+\xi_n}e^{-\mathrm{i}\theta_n},\quad |n|=0,$$

and the Birkhoff normal form (3.4) becomes

$$\tilde H=\big\langle\tilde\omega(\xi)I\big\rangle+\sum_{|n|>0}\tilde\Omega_n(\xi)q_n\bar q_n+\tilde P(\theta,I,q,\bar q,\xi),$$

where $\tilde\omega(\xi)=(\dots,\tilde\omega_n(\xi),\dots)_{|n|=0}$ with

$$\tilde\omega_n(\xi)=n^2+\frac{1}{5}-\frac{1}{2\pi}\xi_n,$$

and

$$\tilde\Omega_n(\xi)=n^2+\frac{1}{5},$$

$$\begin{aligned}\tilde P&=K-\frac{1}{4\pi}\sum_{|n|=0}I_n^2-\frac{1}{4\pi}\sum_{|n|>0}|q_n|^4\\&=-\frac{1}{4\pi}\sum_{|n|=0}I_n^2-\frac{1}{4\pi}\sum_{|n|>0}|q_n|^4\\&\quad+O\big(|\xi|^3+|\xi|^{\frac{5}{2}}|q|+|\xi|^2|q|^2+|\xi|^2|I|+|\xi|^{\frac{3}{2}}|q|^3+|\xi||q|^4+|\xi|^{\frac{1}{2}}|q|^5\big)\end{aligned}$$

with the variables $\{q_n,\bar q_n\}_{|n|=0}$ in $K$ expressed in terms of $I$, $\theta$.

Consider the Taylor–Fourier expansion of $\tilde P$:

$$\tilde P=\sum_{k,\alpha,\beta}\tilde P_{k\alpha\beta}(I)e^{\mathrm{i}\langle k,\theta\rangle}q^\alpha\bar q^\beta.$$

In [15–17], the authors observed that

$$\tilde P_{k\alpha\beta}(I)=0,\quad\text{whenever}\ \sum_{|n|=0}k_n n+\sum_{|n|>0}(\alpha_n-\beta_n)n\neq 0.$$

Such a simple property will be preserved under the KAM iteration.

Now, let $\varepsilon>0$ be sufficiently small. By considering the re-scalings: $\xi\to\varepsilon^8\xi$, $q\to\varepsilon^5 q$, and $I\to\varepsilon^{10}I$, we obtain the re-scaled Hamiltonian

$$\begin{aligned}H(I,\theta,q,\bar q,\xi)&=\varepsilon^{-18}\tilde H\big(\varepsilon^{10}I,\theta,\varepsilon^5 q,\varepsilon^5\bar q,\varepsilon^8\xi\big)\\&=\big\langle\omega(\xi),I\big\rangle+\sum_{|n|>0}\Omega_n(\xi)q_n\bar q_n+P(I,\theta,q,\bar q,\xi,\varepsilon),\end{aligned}\tag{3.5}$$

where $\omega(\xi)=(\dots,\omega_n(\xi),\dots)_{|n|=0}$ with

$$\omega_n(\xi) = \varepsilon^{-8}\left(n^2 + \frac{1}{5}\right) - \frac{1}{2\pi}\xi_n,$$

$$\Omega_n(\xi) = \varepsilon^{-8}\left(n^2 + \frac{1}{5}\right), \quad |n| > 0,$$

$$P = \acute{P} + \grave{P},$$

$$\acute{P} = -\frac{1}{4\pi}\varepsilon^2 \sum_{|n|=0} I_n^2 - \frac{1}{4\pi}\varepsilon^2 \sum_{|n|>0} |q_n|^4,$$

$$\begin{aligned}\grave{P} &= O\left(\varepsilon^6|\xi|^3 + \varepsilon^7|\xi|^{\frac{5}{2}}|q| + \varepsilon^8|\xi|^2|q|^2 + \varepsilon^8|\xi|^2|I| + \varepsilon^9|\xi|^{\frac{3}{2}}|q|^3 + \varepsilon^{10}|\xi||q|^4 + \varepsilon^{11}|\xi|^{\frac{1}{2}}|q|^5\right)\\ &= \varepsilon^6 O\left(|\xi|^3 + \varepsilon|\xi|^{\frac{5}{2}}|q| + \varepsilon^2|\xi|^2|q|^2 + \varepsilon^2|\xi|^2|I| + \varepsilon^3|\xi|^{\frac{3}{2}}|q|^3 + \varepsilon^4|\xi||q|^4 + \varepsilon^5|\xi|^{\frac{1}{2}}|q|^5\right).\end{aligned}$$

**Remark.** In the next KAM iteration, we excite more oscillators in order to increase the number of tangential frequencies, the accompanied problem is that we need more parameters for measure estimates, which will be fulfilled from amplitude-frequency modulation, hence the normal coordinates $q_n, \bar{q}_n$ with $|n| = 1$ will be transformed into action-angle variables. Note that $\frac{1}{4\pi}\varepsilon^2 \sum_{|n|>0} |q_n|^4$ is large compared to $\grave{P}$, hence, the leading term $\frac{1}{4\pi}\varepsilon^2 \sum_{|n|>0} |q_n|^4$ will twist the new added tangential frequencies such that the small divisor conditions still hold true.

Let $\mathcal{O} = [1, 2] \subset \mathbb{R}_+$, $\gamma = \varepsilon^{\frac{3}{4}}$, $\tau = 25$ and set $\Omega = (\bar{\Omega}, \tilde{\Omega})$, $\bar{\Omega} = (\ldots, \Omega_n, \ldots)_{|n|=1}$, $\tilde{\Omega} = (\ldots, \Omega_n, \ldots)_{|n|>1}$. (Note in the next step, $\bar{\Omega}$ will be changed into tangential frequencies in order to increase the number of tangential frequencies.) Moreover, for all integer vectors $(k, j, l) \in \mathbb{Z}^1 \times \mathbb{Z}^2 \times \mathbb{Z}^\infty$, let $|k|$, $|j|$, $|l|$ denote $\ell^1$-norm. We consider the set $\mathcal{O}_0$ consisting of all $\xi \in \mathcal{O}$ such that under the condition $\sum_{|n|=0} k_n n + \sum_{|n|=1} j_n n + \sum_{|n|>1} l_n n = 0$,

$$\left|\left\langle k, \omega(\xi)\right\rangle + \langle j, \bar{\Omega}\rangle + \langle l, \tilde{\Omega}\rangle\right| \geqslant \frac{\gamma}{|k|^\tau}, \quad |k| + |j| + |l| \neq 0,\ |j| + |l| \leqslant 4,\ |l| \leqslant 3.$$

**Proposition 2.** $\operatorname{meas}(\mathcal{O} \setminus \mathcal{O}_0) = O(\varepsilon^{\frac{1}{20}})$.

**Proof.** We first consider the case $k = 0$. Then due to the conditions $0 < |j| + |l| \leqslant 4$, $\sum_{|n|=1} j_n n + \sum_{|n|>1} l_n n = 0$, through elementary calculation, we can get

$$\left|\langle j, \bar{\Omega}\rangle + \langle l, \tilde{\Omega}\rangle\right| \geqslant \frac{1}{5}\varepsilon^{-8}.$$

In fact, during the KAM iteration, because the perturbation is small enough, then the normal frequencies only have small drift, hence the modulated normal frequencies still obey the following estimate

$$\left|\langle j, \bar{\Omega}\rangle + \langle l, \tilde{\Omega}\rangle\right| \geqslant \frac{1}{6}\varepsilon^{-8}.$$

Hence for $k = 0$, the small divisor conditions are automatically satisfied.

We consider the case $k \neq 0$ below. Re-write $\omega(\xi)$, $\bar{\Omega}(\xi)$, $\tilde{\Omega}(\xi)$ as

$$\omega(\xi) = \alpha + A\xi,$$

$$\bar{\Omega}(\xi) = \bar{\beta}, \qquad \tilde{\Omega}(\xi) = \tilde{\beta},$$

where $\alpha = (\ldots, \varepsilon^{-8}(n^2 + \frac{1}{5}), \ldots)_{|n|=0}$, $\bar{\beta} = (\ldots, \varepsilon^{-8}(n^2 + \frac{1}{5}), \ldots)_{|n|=1}$, $\tilde{\beta} = (\ldots, \varepsilon^{-8}(n^2 + \frac{1}{5}), \ldots)_{|n|>1}$, $A = -\frac{1}{2\pi}$.

Since $A=-\frac{1}{2\pi}\neq 0$, then $\langle k,\omega(\xi)\rangle+\langle j,\bar{\Omega}\rangle+\langle l,\tilde{\Omega}\rangle$ are nontrivial affine functions of $\xi\in\mathcal{O}$. The desired measure estimate of $\operatorname{meas}(\mathcal{O}\setminus\mathcal{O}_0)$ then follows from the same argument as that in Section 5.3. □

**Remark.** In the case of Dirichlet boundary conditions, we need to truncate $|k|\leqslant K$ $(K\sim|\ln\varepsilon|)$ and $|\sum_{n=1}\pm j_n n+\sum_{n>1}\pm l_n n|\leqslant K$,

$$\left|\langle k,\omega(\xi)\rangle+\langle j,\bar{\Omega}\rangle+\langle l,\tilde{\Omega}\rangle\right|\geqslant\frac{\gamma}{K^\tau},\quad |k|+|j|+|l|\neq 0,\ |j|+|l|\leqslant 4,\ |l|\leqslant 3,\quad |j|=0,\ \text{if } |l|=3,$$

the above Proposition 2 still holds true (see Section 5.3).

## 4. KAM step

In what follows, we will perform KAM iterations to (3.5) which involves infinite many successive steps, to eliminate lower order $\theta$-dependent terms in $\grave{P}$. Each KAM step and action-angle transformation in the smaller domain (see the following scale corresponding to the smaller domain) will make the perturbation smaller than the previous one at a cost of excluding a small measure set of parameters. At the end, the KAM iterations will be convergent and the measure of the total excluded set will remain small.

More concretely, at each KAM step, firstly we eliminate lower order $\theta$-dependent terms in $\grave{P}$, after that, we perform finite times (76 times in this paper) normal form transformation such that the new perturbation is as small as possible. (The normal form transformation is inessential, one can try to adjust parameters to avoid it.) Then we make the further action-angle coordinate change and get new action-angle variables, thus in a smaller domain (conventionally, people denote the smaller domain by $s_{\nu+1}=\varepsilon_\nu^{\frac{1}{3}}s_\nu=\varepsilon_\nu^{\frac{1}{3}}\varepsilon_{\nu-1}^{\frac{1}{3}}s_{\nu-1}=\cdots=\prod_{j=0}^{\nu}\varepsilon_j^{\frac{1}{3}}s_0$, however in this paper, the scaling techniques are adopted to unify action-angle coordinate change and shrinking domain), we get a new Hamiltonian with new perturbation smaller than the previous one.

To begin with the KAM iteration, we fix $r,s,\rho>0$ and restrict the Hamiltonian (3.5) to the domain $D(r,s)$ and restrict the parameter to the set $\mathcal{O}_0$. Initially, we set $\omega^0=\omega$, $\xi(0)=\xi_0=\xi$, $I(0)=I=(\dots,I_n,\dots)_{|n|=0}$, $\theta(0)=\theta=(\dots,\theta_n,\dots)_{|n|=0}$, $\Omega_n^0=\Omega_n$, $P_0=P$, $r_0=r$, $\gamma_0=\gamma$, $\tau_0=\tau$, and

$$N_0=\left\langle\omega^0\big(\xi(0)\big),I(0)\right\rangle+\sum_{|n|>0}\Omega_n^0\big(\xi(0)\big)q_n\bar{q}_n,$$

$$H_0=N_0+P_0,$$

$$P_0=\acute{P}_0+\grave{P}_0,$$

$$\acute{P}_0=-\frac{1}{4\pi}\varepsilon_0^2\sum_{|n|=0}I_n^2-\frac{1}{4\pi}\varepsilon_0^2\sum_{|n|>0}|q_n|^4,$$

$$\grave{P}_0=O\big(\varepsilon_0^6\big).$$

Hence, $H_0$ is real analytic on $D(r_0,s)$ and also depends on $\xi(0)\in\mathcal{O}_0$ smoothly in the sense of Whitney, and

$$\|X_{\acute{P}_0}\|_{D(r_0,s),\mathcal{O}_0}\leqslant\varepsilon^2\equiv\varepsilon_0^2,\qquad \|X_{\grave{P}_0}\|_{D(r_0,s),\mathcal{O}_0}\leqslant\varepsilon^6\equiv\varepsilon_0^6.$$

We recall that, under the condition $\sum_{|n|=0}k_n n+\sum_{|n|=1}j_n n+\sum_{|n|>1}l_n n=0$,

$$\mathcal{O}_0=\left\{\xi(0):\ \left|\left\langle k,\omega^0\big(\xi(0)\big)\right\rangle+\left\langle j,\bar{\Omega}^0\right\rangle+\left\langle l,\tilde{\Omega}^0\right\rangle\right|\geqslant\frac{\gamma_0}{|k|^{\tau_0}},\ \begin{matrix}|k|+|j|+|l|\neq 0,\\ |j|+|l|\leqslant 4,\ |l|\leqslant 3\end{matrix}\right\},$$

and $P_0 = \sum_{k,\alpha,\beta} P_{0k\alpha\beta}(I(0))e^{\mathrm{i}\langle k,\theta(0)\rangle} q^\alpha \bar{q}^\beta$ has a compact form, i.e.,

$$P_{0k\alpha\beta} = 0, \quad \text{whenever } \sum_{|n|=0} k_n n + \sum_{|n|>0} (\alpha_n - \beta_n) n \neq 0.$$

After the first KAM step, we arrive at the Hamiltonian

$$\begin{gathered}
H_1 = N_1 + P_1, \\
N_1 = \bigl\langle \omega^1\bigl(\xi(0), \xi(1)\bigr), I(1)\bigr\rangle + \sum_{|n|>1} \Omega_n^1\bigl(\xi(0), \xi(1)\bigr) q_n \bar{q}_n, \\
\xi(0) = \xi_0, \qquad \xi(1) = (\xi_1, \xi_{-1}), \\
\omega^1\bigl(\xi(0), \xi(1)\bigr) = \bigl(\omega_0\bigl(\xi(0), \xi(1)\bigr), \omega_1\bigl(\xi(0), \xi(1)\bigr), \omega_{-1}\bigl(\xi(0), \xi(1)\bigr)\bigr), \\
\begin{aligned}
\omega_0\bigl(\xi(0), \xi(1)\bigr) = {} & \varepsilon_0^{-8} \varepsilon_1^{-8}\left(0^2 + \frac{1}{5}\right) + \left(-\frac{1}{2\pi}\right)\varepsilon_1^{-8}\xi_0 \\
& + O\bigl(\varepsilon_0^6 \varepsilon_1^{-8}\bigr) f_0^1\bigl(\xi(0)\bigr) + O\bigl(\varepsilon_0^6\bigr) f_1^1\bigl(\xi(1)\bigr), \\
\omega_1\bigl(\xi(0), \xi(1)\bigr) = {} & \varepsilon_0^{-8} \varepsilon_1^{-8}\left(1^2 + \frac{1}{5}\right) + \left(-\frac{1}{2\pi}\right)\varepsilon_0^2\xi_1 \\
& + O\bigl(\varepsilon_0^6 \varepsilon_1^{-8}\bigr) f_0^1\bigl(\xi(0)\bigr) + O\bigl(\varepsilon_0^6\bigr) f_1^1\bigl(\xi(1)\bigr), \\
\omega_{-1}\bigl(\xi(0), \xi(1)\bigr) = {} & \varepsilon_0^{-8} \varepsilon_1^{-8}\left((-1)^2 + \frac{1}{5}\right) + \left(-\frac{1}{2\pi}\right)\varepsilon_0^2\xi_{-1} \\
& + O\bigl(\varepsilon_0^6 \varepsilon_1^{-8}\bigr) f_0^1\bigl(\xi(0)\bigr) + O\bigl(\varepsilon_0^6\bigr) f_1^1\bigl(\xi(1)\bigr),
\end{aligned} \\
\bigl\| f_0^1\bigl(\xi(0)\bigr)\bigr\|_{\mathcal{O}_1} = \sup_{\mathcal{O}_1}\left(\bigl|f_0^1\bigl(\xi(0)\bigr)\bigr| + \left|\frac{\partial f_0^1}{\partial \xi(0)}\right|\right) \leqslant 1, \\
\bigl\| f_1^1\bigl(\xi(1)\bigr)\bigr\|_{\mathcal{O}_1} = \sup_{\mathcal{O}_1}\left(\bigl|f_1^1\bigl(\xi(1)\bigr)\bigr| + \left|\frac{\partial f_1^1}{\partial \xi(1)}\right|\right) \leqslant 1, \\
I(1) = (\ldots, I_n, \ldots)_{|n|\leqslant 1}, \qquad \theta(1) = (\ldots, \theta_n, \ldots)_{|n|\leqslant 1}, \\
\begin{aligned}
\Omega_n^1\bigl(\xi(0), \xi(1)\bigr) = {} & \varepsilon_0^{-8}\varepsilon_1^{-8}\left(n^2 + \frac{1}{5}\right) \\
& + O\bigl(\varepsilon_0^6 \varepsilon_1^{-8}\bigr) f_0^1\bigl(\xi(0)\bigr) + O\bigl(\varepsilon_0^6\bigr) f_1^1\bigl(\xi(1)\bigr), \quad |n| > 1,
\end{aligned} \\
P_1 = \acute{P}_1 + \grave{P}_1, \\
\acute{P}_1 = -\frac{1}{4\pi}\left(\prod_{i=0}^{1}\varepsilon_i^2\right)\sum_{|n|\leqslant 1} I_n^2 - \frac{1}{4\pi}\left(\prod_{i=0}^{1}\varepsilon_i^2\right)\sum_{|n|>1} |q_n|^4, \\
\grave{P}_1 = O\bigl(\varepsilon_0^6\varepsilon_1^2\bigr),
\end{gathered}$$

which is real analytic in $(\theta(1), I(1), q, \bar{q}) \in D_1 = D(r_1, s)$, and depends on $(\xi(0), \xi(1)) \in \mathcal{O}_1 \subset [1, 2]^3$ Whitney smoothly, where under the condition

$$\sum_{|n|\leqslant 1} k_n n + \sum_{|n|=2} j_n n + \sum_{|n|>2} l_n n = 0,$$

$$\mathcal{O}_1 = \left\{ \big(\xi(0), \xi(1)\big):\ \big|\langle k, \omega^1\rangle + \langle j, \bar{\Omega}^1\rangle + \langle l, \tilde{\Omega}^1\rangle\big| \geqslant \frac{\gamma_1}{|k|^{\tau_1}},\ \begin{array}{l} |k|+|j|+|l| \neq 0, \\ |j|+|l| \leqslant 4,\ |l| \leqslant 3 \end{array} \right\},$$

for $\gamma_1 = \varepsilon_0 \varepsilon_1^{\frac{3}{4}}$, $\tau_1 = 50$, $\varepsilon_1 = \varepsilon_0^{\frac{9}{5}}$ and $\bar{\Omega}^1 = (\ldots, \Omega_n^1, \ldots)_{|n|=2}$, $\tilde{\Omega}^1 = (\ldots, \Omega_n^1, \ldots)_{|n|>2}$. We also assume that

$$\|X_{\acute{P}_1}\|_{D_1,\mathcal{O}_1} \leqslant \prod_{i=0}^{1} \varepsilon_i^2, \qquad \|X_{\grave{P}_1}\|_{D_1,\mathcal{O}_1} \leqslant \varepsilon_0^6 \varepsilon_1^2,$$

and that $P_1 = \sum_{k,\alpha,\beta} P_{1k\alpha\beta}(I(1)) e^{\mathrm{i}\langle k, \theta(1)\rangle} q^\alpha \bar{q}^\beta$ has a compact form, i.e.,

$$P_{1k\alpha\beta} = 0, \quad \text{whenever} \sum_{|n|\leqslant 1} k_n n + \sum_{|n|>1} (\alpha_n - \beta_n) n \neq 0.$$

Suppose that after a $\nu$th KAM step, we arrive at a Hamiltonian

$$H_\nu = N_\nu + P_\nu,$$

$$N_\nu = \big\langle \omega^\nu\big(\xi(0), \ldots, \xi(\nu)\big), I(\nu)\big\rangle + \sum_{|n|>\nu} \Omega_n^\nu\big(\xi(0), \ldots, \xi(\nu)\big) q_n \bar{q}_n,$$

$$\xi(i) = (\ldots, \xi_n, \ldots)_{|n|=i}, \quad 0 \leqslant i \leqslant \nu,$$

$$\omega^\nu\big(\xi(0), \ldots, \xi(\nu)\big) = \big(\ldots, \omega_n\big(\xi(0), \ldots, \xi(\nu)\big), \ldots\big)_{|n|\leqslant \nu},$$

$$\begin{aligned} \omega_n\big(\xi(0), \ldots, \xi(\nu)\big) &= \varepsilon_0^{-8} \cdots \varepsilon_\nu^{-8} \left(n^2 + \frac{1}{5}\right) + \left(-\frac{1}{2\pi}\right) \varepsilon_0^2 \varepsilon_1^2 \cdots \varepsilon_{i-1}^2 \varepsilon_{i+1}^{-8} \cdots \varepsilon_\nu^{-8} \xi_n \\ &\quad + O\big(\varepsilon_0^6 \varepsilon_1^{-8} \cdots \varepsilon_\nu^{-8}\big) f_0^\nu\big(\xi(0)\big) + O\big(\varepsilon_0^6 \varepsilon_2^{-8} \cdots \varepsilon_\nu^{-8}\big) f_1^\nu\big(\xi(1)\big) \\ &\quad + \cdots + O\big(\varepsilon_0^6 \varepsilon_1^2 \cdots \varepsilon_{i-1}^2 \varepsilon_{i+1}^{-8} \cdots \varepsilon_\nu^{-8}\big) f_i^\nu\big(\xi(i)\big) \\ &\quad + \cdots + O\big(\varepsilon_0^6 \varepsilon_1^2 \cdots \varepsilon_{\nu-1}^2\big) f_\nu^\nu\big(\xi(\nu)\big), \quad |n| = i, \end{aligned}$$

$$\big\| f_i^\nu\big(\xi(i)\big)\big\|_{\mathcal{O}_\nu} = \sup_{\mathcal{O}_\nu} \left( \big|f_i^\nu\big(\xi(i)\big)\big| + \left|\frac{\partial f_i^\nu}{\partial \xi(i)}\right| \right) \leqslant 1, \quad 0 \leqslant i \leqslant \nu,$$

$$I(\nu) = (\ldots, I_n, \ldots)_{|n|\leqslant \nu}, \qquad \theta(\nu) = (\ldots, \theta_n, \ldots)_{|n|\leqslant \nu},$$

$$\begin{aligned} \Omega_n^\nu\big(\xi(0), \ldots, \xi(\nu)\big) &= \varepsilon_0^{-8} \cdots \varepsilon_\nu^{-8} \left(n^2 + \frac{1}{5}\right) \\ &\quad + O\big(\varepsilon_0^6 \varepsilon_1^{-8} \cdots \varepsilon_\nu^{-8}\big) f_0^\nu\big(\xi(0)\big) + O\big(\varepsilon_0^6 \varepsilon_2^{-8} \cdots \varepsilon_\nu^{-8}\big) f_1^\nu\big(\xi(1)\big) \\ &\quad + \cdots + O\big(\varepsilon_0^6 \varepsilon_1^2 \cdots \varepsilon_{i-1}^2 \varepsilon_{i+1}^{-8} \cdots \varepsilon_\nu^{-8}\big) f_i^\nu\big(\xi(i)\big) \\ &\quad + \cdots + O\big(\varepsilon_0^6 \varepsilon_1^2 \cdots \varepsilon_{\nu-1}^2\big) f_\nu^\nu\big(\xi(\nu)\big), \quad |n| > \nu, \end{aligned}$$

$$P_\nu = \acute{P}_\nu + \grave{P}_\nu,$$

$$\acute{P}_\nu = -\frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|\leqslant\nu} I_n^2 - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|>\nu}|q_n|^4,$$

$$\grave{P}_\nu = O\left(\varepsilon_0^6\left(\prod_{i=1}^{\nu}\varepsilon_i^2\right)\right),$$

which is real analytic in $(\theta(\nu), I(\nu), q, \bar{q}) \in D_\nu = D(r_\nu, s)$, and depends on $(\xi(0), \ldots, \xi(\nu)) \in \mathcal{O}_\nu \subset [1,2]^{2\nu+1}$ Whitney smoothly, where under the condition

$$\sum_{|n|\leqslant\nu} k_n n + \sum_{|n|=\nu+1} j_n n + \sum_{|n|>\nu+1} l_n n = 0,$$

$$\mathcal{O}_\nu = \left\{ \big(\xi(0), \ldots, \xi(\nu)\big)\colon \left|\left\langle k, \omega^\nu\right\rangle + \left\langle j, \bar{\Omega}^\nu\right\rangle + \left\langle l, \tilde{\Omega}^\nu\right\rangle\right| \geqslant \frac{\gamma_\nu}{|k|^{\tau_\nu}},\ \begin{array}{l} |k|+|j|+|l|\neq 0, \\ |j|+|l|\leqslant 4,\ |l|\leqslant 3 \end{array} \right\},$$

for $\gamma_\nu = (\prod_{i=0}^{\nu-1}\varepsilon_i)\varepsilon_\nu^{\frac{3}{4}}$, $\tau_\nu = 25(\nu+1)$, $\varepsilon_\nu = \varepsilon_{\nu-1}^{\frac{9}{5}}$ and $\bar{\Omega}^\nu = (\ldots, \Omega_n^\nu, \ldots)_{|n|=\nu+1}$, $\tilde{\Omega}^\nu = (\ldots, \Omega_n^\nu, \ldots)_{|n|>\nu+1}$. We also assume that

$$\|X_{\acute{P}_\nu}\|_{D_\nu, \mathcal{O}_\nu} \leqslant \prod_{i=0}^{\nu}\varepsilon_i^2, \qquad \|X_{\grave{P}_\nu}\|_{D_\nu, \mathcal{O}_\nu} \leqslant \varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2,$$

and that $P_\nu = \sum_{k,\alpha,\beta} P_{\nu k\alpha\beta}(I(\nu))e^{\mathrm{i}\langle k,\theta(\nu)\rangle}q^\alpha\bar{q}^\beta$ has a compact form, i.e.,

$$P_{\nu k\alpha\beta} = 0, \quad \text{whenever} \sum_{|n|\leqslant\nu} k_n n + \sum_{|n|>\nu}(\alpha_n - \beta_n)n \neq 0.$$

We will construct a symplectic transformation $\Phi = \Phi_\nu$, which, in smaller frequency and phase domains, carries the above Hamiltonian into the next KAM cycle.

**Remark.** In the case of Dirichlet boundary conditions, we need

$$|k| \leqslant K_\nu, \qquad \left|\sum_{n=\nu+1} \pm j_n n + \sum_{n>\nu+1} \pm l_n n\right| \leqslant K_\nu,$$

$$\mathcal{O}_\nu = \left\{ \big(\xi(0), \ldots, \xi(\nu)\big)\colon \left|\left\langle k, \omega^\nu\right\rangle + \left\langle j, \bar{\Omega}^\nu\right\rangle + \left\langle l, \tilde{\Omega}^\nu\right\rangle\right| \geqslant \frac{\gamma_\nu}{K_\nu^{\tau_\nu}},\ \begin{array}{l} |k|+|j|+|l|\neq 0, \\ |j|+|l|\leqslant 4,\ |l|\leqslant 3, \\ |j|=0,\ \text{if } |l|=3 \end{array} \right\}.$$

*4.1. Truncation*

We expand $\grave{P}_\nu$ into the Fourier–Taylor series

$$\grave{P}_\nu = \sum_{k,l,\alpha,\beta} \grave{P}_{\nu kl\alpha\beta} e^{\mathrm{i}\langle k,\theta(\nu)\rangle} I(\nu)^l q^\alpha \bar{q}^\beta,$$

where $k \in \mathbb{Z}^{2\nu+1}$, $l \in \mathbb{N}^{2\nu+1}$ and $\alpha = (\ldots, \alpha_n, \ldots)_{|n|>\nu}$, $\beta = (\ldots, \beta_n, \ldots)_{|n|>\nu}$, $\alpha_n, \beta_n \in \mathbb{N}$, are multi-indices with finitely many non-vanishing components. For the convenience of notations, we denote

$$\bar\alpha=(\ldots,\alpha_n,\ldots)_{|n|=\nu+1},\qquad \tilde\alpha=(\ldots,\alpha_n,\ldots)_{|n|>\nu+1},$$
$$\bar\beta=(\ldots,\beta_n,\ldots)_{|n|=\nu+1},\qquad \tilde\beta=(\ldots,\beta_n,\ldots)_{|n|>\nu+1},$$

and correspondingly, $z=(\ldots,q_n,\ldots)_{|n|=\nu+1}$, abusing the notations, we still denote $q=(\ldots,q_n,\ldots)_{|n|>\nu+1}$.

Let $R_\nu$ be the following truncation of $\grave{P}_\nu$:

$$R_\nu\big(\theta(\nu),I(\nu),z,\bar z,q,\bar q\big)=\sum_{\substack{2|l|+|\bar\alpha+\bar\beta|\leqslant 4\\ 2|l|+|\tilde\alpha+\tilde\beta|\leqslant 3\\ |\bar\alpha+\bar\beta|+|\tilde\alpha+\tilde\beta|\leqslant 4}}\sum_k \grave{P}_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}e^{\mathrm{i}\langle k,\theta(\nu)\rangle}I(\nu)^l z^{\bar\alpha}\bar z^{\bar\beta}q^{\tilde\alpha}\bar q^{\tilde\beta}.$$

The mean value of $R_\nu$ is defined as

$$\begin{aligned}[]
[R_\nu]&=\sum_{\substack{|l|+|\bar\alpha|\leqslant 2\\ |l|+|\tilde\alpha|\leqslant 1\\ |\bar\alpha|+|\tilde\alpha|\leqslant 2}}\grave{P}_{\nu 0l\bar\alpha\bar\alpha\tilde\alpha\tilde\alpha}I(\nu)^l z^{\bar\alpha}\bar z^{\bar\alpha}q^{\tilde\alpha}\bar q^{\tilde\alpha}\\
&=\sum_{|l|\leqslant 1}\grave{P}_{\nu 0l0000}I(\nu)^l+\sum_{|n|=\nu+1}\grave{P}^{011}_{\nu nn}q_n\bar q_n\\
&\quad+\sum_{|m|=|n|=\nu+1}\grave{P}^{01111}_{\nu mmnn}q_m\bar q_m q_n\bar q_n+\sum_{|l|=1,\,|n|=\nu+1}\grave{P}^{011}_{\nu lnn}q_n\bar q_n I(\nu)^l\\
&\quad+\sum_{|n|>\nu+1}\grave{P}^{011}_{\nu nn}q_n\bar q_n+\sum_{|m|=\nu+1,\,|n|>\nu+1}\grave{P}^{01111}_{\nu mmnn}q_m\bar q_m q_n\bar q_n.
\end{aligned}$$

Note that $\grave{P}_\nu$ has a compact form,

$$\grave{P}_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}=0,\quad \text{if}\ \sum_{|n|\leqslant\nu}k_n n+\sum_{|n|=\nu+1}(\bar\alpha_n-\bar\beta_n)n+\sum_{|n|>\nu+1}(\tilde\alpha_n-\tilde\beta_n)n\neq 0.$$

By definition of the weighted norms, we clearly have

$$\|X_{R_\nu}\|_{D(r_\nu,s),\mathcal{O}_\nu}\leqslant\|X_{\grave{P}_\nu}\|_{D(r_\nu,s),\mathcal{O}_\nu}\leqslant\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2.$$

**Remark.** In the case of Dirichlet boundary conditions, we need further truncate $R_\nu$ with $|k|\leqslant K_\nu$ and $|\sum_{n=\nu+1}\pm(\bar\alpha_n\pm\bar\beta_n)n+\sum_{n>\nu+1}\pm(\tilde\alpha_n\pm\tilde\beta_n)n|\leqslant K_\nu$ and $|\bar\alpha+\bar\beta|=0$ whenever $|\tilde\alpha+\tilde\beta|=3$.

### *4.2. The homological equation*

Let $r_{\nu+1}=\frac{r_\nu}{2}+\frac{r_0}{4}$. We now look for a real analytic function $F_\nu$, defined in the smaller domain $D(r_{\nu+1},s)$ such that the time-1 map $\Phi=\Phi^1_{F_\nu}:D(r_{\nu+1},s)\to D(r_\nu,s)$ of the Hamiltonian flow $\Phi^t_{F_\nu}$ associated with $F_\nu$ transforms $H_\nu$ into the Hamiltonian $H_{\nu+1}$ in the next KAM cycle. Let $F_\nu$ have the form

$$F_\nu\big(\theta(\nu),I(\nu),z,\bar z,q,\bar q\big)=\sum_{\substack{2|l|+|\bar\alpha+\bar\beta|\leqslant 4\\ 2|l|+|\tilde\alpha+\tilde\beta|\leqslant 3\\ |\bar\alpha+\bar\beta|+|\tilde\alpha+\tilde\beta|\leqslant 4}}\sum_{\substack{k\\ |k|+|\bar\alpha-\bar\beta|+|\tilde\alpha-\tilde\beta|\neq 0}}F_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}e^{\mathrm{i}\langle k.\theta(\nu)\rangle}I(\nu)^l z^{\bar\alpha}\bar z^{\bar\beta}q^{\tilde\alpha}\bar q^{\tilde\beta},$$

which satisfies a compact form,

$$F_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}=0,\quad \text{if} \sum_{|n|\leqslant\nu} k_n n+\sum_{|n|=\nu+1}(\bar\alpha_n-\bar\beta_n)n+\sum_{|n|>\nu+1}(\tilde\alpha_n-\tilde\beta_n)n\neq 0$$

and the homological equation

$$\{N_\nu,F_\nu\}+R_\nu-[R_\nu]=0. \tag{4.1}$$

**Remark.** In the case of Dirichlet boundary conditions, we need further truncate $F_\nu$ with $|k|\leqslant K_\nu$ and $|\sum_{n=\nu+1}\pm(\bar\alpha_n\pm\bar\beta_n)n+\sum_{n>\nu+1}\pm(\tilde\alpha_n\pm\tilde\beta_n)n|\leqslant K_\nu$ and $|\bar\alpha+\bar\beta|=0$ whenever $|\tilde\alpha+\tilde\beta|=3$, such that $F_\nu$ has the same form as $R_\nu$.

By comparing coefficients, it is easy to see that the homological equation (4.1) is equivalent to

$$\big(\langle k,\omega^\nu\rangle-\langle\bar\alpha-\bar\beta,\bar\Omega^\nu\rangle-\langle\tilde\alpha-\tilde\beta,\tilde\Omega^\nu\rangle\big)F_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}=\mathrm{i}\grave{P}_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}. \tag{4.2}$$

Hence the homological equation (4.1) is uniquely solvable on $\mathcal{O}_\nu$ to yield the function $F_\nu$ which is real analytic in $(\theta(\nu),I(\nu),z,\bar z,q,\bar q)$ and Whitney smooth in $(\xi(0),\dots,\xi(\nu))\in\mathcal{O}_\nu$.

**Lemma 4.1.** *Let $D_3=D(r_{\nu+1}+\frac{3}{4}(r_\nu-r_{\nu+1}),s)$, then*

$$\|X_{F_\nu}\|_{D_3,\mathcal{O}_\nu}\leqslant\gamma_\nu^{-2}\left(\frac{8\tau_\nu+8}{(r_\nu-r_{\nu+1})e}\right)^{2\tau_\nu+2}\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2. \tag{4.3}$$

**Proof.** By (4.2), the definition of $\mathcal{O}_\nu$ and the definition of the weighted norms, we have

$$|F_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}|_{\mathcal{O}_\nu}\leqslant\gamma_\nu^{-2}|k|^{2\tau_\nu+1}|\grave{P}_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}|_{\mathcal{O}_\nu}. \tag{4.4}$$

It follows that

$$\begin{aligned}
&\frac{1}{s^2}\|F_{\nu_{\theta(\nu)}}\|_{D_3,\mathcal{O}_\nu}\\
&\leqslant\frac{1}{s^2}\sum_{\substack{2|l|+|\bar\alpha+\bar\beta|\leqslant 4\\ 2|l|+|\tilde\alpha+\tilde\beta|\leqslant 3\\ |\bar\alpha+\bar\beta|+|\tilde\alpha+\tilde\beta|\leqslant 4}}\ \sum_{\substack{k\\ |k|+|\bar\alpha-\bar\beta|+|\tilde\alpha-\tilde\beta|\neq 0}}|F_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}|_{\mathcal{O}_\nu}|k|e^{|k|(r_\nu-\frac14(r_\nu-r_{\nu+1}))}\big|I(\nu)\big|^l|z|^{\bar\alpha}|\bar z|^{\bar\beta}|q|^{\tilde\alpha}|\bar q|^{\tilde\beta}\\
&\leqslant\frac{1}{s^2}\sum_{\substack{2|l|+|\bar\alpha+\bar\beta|\leqslant 4\\ 2|l|+|\tilde\alpha+\tilde\beta|\leqslant 3\\ |\bar\alpha+\bar\beta|+|\tilde\alpha+\tilde\beta|\leqslant 4}}\ \sum_k\gamma_\nu^{-2}|k|^{2\tau_\nu+2}e^{-\frac14|k|(r_\nu-r_{\nu+1})}|\grave{P}_{\nu kl\bar\alpha\bar\beta\tilde\alpha\tilde\beta}|_{\mathcal{O}_\nu}e^{|k|r_\nu}\big|I(\nu)\big|^l|z|^{\bar\alpha}|\bar z|^{\bar\beta}|q|^{\tilde\alpha}|\bar q|^{\tilde\beta}\\
&\leqslant\gamma_\nu^{-2}\left(\frac{8\tau_\nu+8}{(r_\nu-r_{\nu+1})e}\right)^{2\tau_\nu+2}\|X_{R_\nu}\|\\
&\leqslant\gamma_\nu^{-2}\left(\frac{8\tau_\nu+8}{(r_\nu-r_{\nu+1})e}\right)^{2\tau_\nu+2}\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2.
\end{aligned} \tag{4.5}$$

Similarly,

$$\|F_{\nu_{I(\nu)}}\|_{D_3,\mathcal{O}_\nu} \leqslant \gamma_\nu^{-2}\left(\frac{8\tau_\nu+8}{(r_\nu-r_{\nu+1})e}\right)^{2\tau_\nu+2}\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2,$$

$$\frac{1}{s}\sum_{|n|=\nu+1}\|F_{\nu_{z_n}}\|_{D_3,\mathcal{O}_\nu}e^{|n|\rho} \leqslant \gamma_\nu^{-2}\left(\frac{8\tau_\nu+8}{(r_\nu-r_{\nu+1})e}\right)^{2\tau_\nu+2}\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2,$$

$$\frac{1}{s}\sum_{|n|=\nu+1}\|F_{\nu_{\bar z_n}}\|_{D_3,\mathcal{O}_\nu}e^{|n|\rho} \leqslant \gamma_\nu^{-2}\left(\frac{8\tau_\nu+8}{(r_\nu-r_{\nu+1})e}\right)^{2\tau_\nu+2}\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2,$$

$$\frac{1}{s}\sum_{|n|>\nu+1}\|F_{\nu_{q_n}}\|_{D_3,\mathcal{O}_\nu}e^{|n|\rho} \leqslant \gamma_\nu^{-2}\left(\frac{8\tau_\nu+8}{(r_\nu-r_{\nu+1})e}\right)^{2\tau_\nu+2}\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2,$$

$$\frac{1}{s}\sum_{|n|>\nu+1}\|F_{\nu_{\bar q_n}}\|_{D_3,\mathcal{O}_\nu}e^{|n|\rho} \leqslant \gamma_\nu^{-2}\left(\frac{8\tau_\nu+8}{(r_\nu-r_{\nu+1})e}\right)^{2\tau_\nu+2}\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2.$$

Hence we have

$$\|X_{F_\nu}\|_{D_3,\mathcal{O}_\nu} \leqslant \gamma_\nu^{-2}\left(\frac{8\tau_\nu+8}{(r_\nu-r_{\nu+1})e}\right)^{2\tau_\nu+2}\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2. \qquad \square \tag{4.6}$$

According to our choice of parameters $\gamma_\nu = (\prod_{i=0}^{\nu-1}\varepsilon_i)\varepsilon_\nu^{\frac{3}{4}}$, $\tau_\nu = 25(\nu+1)$, $r_{\nu+1} = \frac{r_\nu}{2} + \frac{r_0}{4}$, $\varepsilon_\nu = \varepsilon_{\nu-1}^{\frac{9}{5}}$, we can get

$$\|X_{F_\nu}\|_{D_3,\mathcal{O}_\nu} \leqslant \varepsilon_\nu^{\frac{1}{2}}. \tag{4.7}$$

In the next lemma, we give some estimates for $\Phi_F^t$. The formula (4.8) will be used to prove our coordinate transformation is well defined. Inequality (4.9) will be used to check the convergence of the iteration.

**Lemma 4.2.** *Let $D_i = D(r_{\nu+1} + \frac{i}{4}(r_\nu - r_{\nu+1}), s)$, $0 < i \leqslant 4$, then we have*

$$\Phi_{F_\nu}^t : D_2 \to D_3, \quad -1 \leqslant t \leqslant 1. \tag{4.8}$$

*Moreover,*

$$\left\|D\Phi_{F_\nu}^t - \mathrm{Id}\right\|_{D_1} < \varepsilon_\nu^{\frac{1}{2}}. \tag{4.9}$$

**Proof.** Let

$$\left\|D^m F\right\|_{D,\mathcal{O}} = \max\left\{\left\|\frac{\partial^{|i|+|l|+|\bar\alpha|+|\bar\beta|+|\tilde\alpha|+|\tilde\beta|}}{\partial\theta^i\partial I^l\partial z^{\bar\alpha}\partial\bar z^{\bar\beta}\partial q^{\tilde\alpha}\partial\bar q^{\tilde\beta}}F\right\|_{D,\mathcal{O}},\ |i|+|l|+|\bar\alpha|+|\bar\beta|+|\tilde\alpha|+|\tilde\beta| = m \geqslant 2\right\}.$$

Notice that $F_\nu$ is a polynomial of degree 1 in $I$ and degree 4 in $z$, $\bar z$ and degree 3 in $q$, $\bar q$. From (4.7), the definition of the weighted norms and the Cauchy inequality, it follows that

$$\left\|D^m F_\nu\right\|_{D_2,\mathcal{O}_\nu} < \varepsilon_\nu^{\frac{1}{2}}, \tag{4.10}$$

for any $m \geqslant 2$.

To get the estimates for $\Phi^t_{F_\nu}$, we start from the integral equation,

$$\Phi^t_{F_\nu} = \mathrm{id} + \int_0^t X_{F_\nu} \circ \Phi^s_{F_\nu}\, ds$$

so that $\Phi^t_{F_\nu} : D_2 \to D_3$, $-1 \leqslant t \leqslant 1$, which follows directly from (4.6). Since

$$D\Phi^t_{F_\nu} = \mathrm{Id} + \int_0^t (DX_{F_\nu}) D\Phi^s_{F_\nu}\, ds = \mathrm{Id} + \int_0^t J\big(D^2 F_\nu\big) D\Phi^s_{F_\nu}\, ds,$$

where $J$ denotes the standard symplectic matrix $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$, it follows that

$$\big\| D\Phi^t_{F_\nu} - \mathrm{Id} \big\|_{D_1} \leqslant 2 \big\| D^2 F_\nu \big\|_{D_2} < \varepsilon_\nu^{\frac{1}{2}}. \tag{4.11}$$

Consequently Lemma 4.2 follows. □

**Remark.** Due to technical reasons, at each KAM step, we have to repeat the above procedure for many times. More precisely, we need $\varepsilon_{\nu+1} = \varepsilon_\nu^{\frac{9}{5}}$, and the new Hamiltonian will be scaled by $\varepsilon^{18}_{\nu+1}$, and after re-scaling, by mathematical induction, the new perturbation will have more small factor $\varepsilon_\nu^2 \varepsilon_{\nu+1}^2$, then we have

$$\varepsilon_\nu^2 \varepsilon_{\nu+1}^2 \varepsilon_{\nu+1}^{18} = \varepsilon_\nu^{38} = \big(\varepsilon_\nu^{\frac{1}{2}}\big)^{76},$$

thus we repeat the above procedure for 76 times at each KAM step, so that our new perturbation has the following form

$$\begin{aligned}
P_{\nu+1} &= \acute{P}_{\nu+1} + \grave{P}_{\nu+1}, \\
\acute{P}_{\nu+1} &= -\frac{1}{4\pi}\left(\prod_{i=0}^{\nu} \varepsilon_i^2\right) \sum_{|n| \leqslant \nu} I_n^2 - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu} \varepsilon_i^2\right) \sum_{|n| = \nu+1} |z_n|^4 - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu} \varepsilon_i^2\right) \sum_{|n| > \nu+1} |q_n|^4, \\
\grave{P}_{\nu+1} &= O\left(\varepsilon_0^6 \left(\prod_{i=1}^{\nu-1} \varepsilon_i^2\right) \varepsilon_\nu^{38}\right) \quad \text{(lower order terms in } I, z, q) \\
&\quad + O\left(\varepsilon_0^6 \prod_{i=1}^{\nu} \varepsilon_i^2\right) \big(|I|^2 + |z|^5 + |I||z|^3 + |z|^3|q|^2 + |I||q|^2 + |q|^4\big).
\end{aligned} \tag{4.12}$$

**Remark.** In the case of Dirichlet boundary conditions, $\grave{P}_{\nu+1}$ should include one more term $O(\varepsilon_0^6 \prod_{i=1}^{\nu} \varepsilon_i^2)|z||q|^3$.

Thus after the further action-angle transformation in the smaller domain (its scale is roughly $\varepsilon_{\nu+1}$, as a consequence, our amplitudes decay super-exponentially such as $\xi_n \sim \varepsilon_0^{(\frac{9}{5})^{|n|}}$), we can assure that the new perturbation is much smaller.

Now let $\Phi_\nu = \Phi^1_{F_\nu}$, $D_{\nu+1} = D(r_{\nu+1}, s)$, then

$$\begin{aligned}
H_{\nu+1} &= N_\nu + [R_\nu] + P_{\nu+1} \\
&= N_\nu + \sum_{\substack{|l|+|\bar\alpha|\leqslant 2\\ |l|+|\tilde\alpha|\leqslant 1\\ |\bar\alpha|+|\tilde\alpha|\leqslant 2}} \grave{P}_{\nu 0 l\bar\alpha\bar\alpha\tilde\alpha\tilde\alpha} I(\nu)^l z^{\bar\alpha}\bar z^{\bar\alpha} q^{\tilde\alpha}\bar q^{\tilde\alpha} + \acute{P}_{\nu+1} + \grave{P}_{\nu+1} \\
&= \big\langle \omega^\nu\big(\xi(0),\dots,\xi(\nu)\big), I(\nu)\big\rangle + \sum_{|n|=\nu+1} \Omega_n^\nu\big(\xi(0),\dots,\xi(\nu)\big) z_n\bar z_n \\
&\quad + \sum_{|n|>\nu+1} \Omega_n^\nu\big(\xi(0),\dots,\xi(\nu)\big) q_n\bar q_n \\
&\quad + \sum_{|l|=1} O\big(\grave{P}_{\nu 0 l 0000}\big) I(\nu)^l + \sum_{|n|=\nu+1} O\big(\grave{P}^{011}_{\nu nn}\big) z_n\bar z_n \\
&\quad + \sum_{|m|=|n|=\nu+1} O\big(\grave{P}^{01111}_{\nu mmnn}\big) z_m\bar z_m z_n\bar z_n \\
&\quad + \sum_{|l|=1,\,|n|=\nu+1} O\big(\grave{P}^{011}_{\nu lnn}\big) z_n\bar z_n I(\nu)^l + \sum_{|n|>\nu+1} O\big(\grave{P}^{011}_{\nu nn}\big) q_n\bar q_n \\
&\quad + \sum_{|m|=\nu+1,\,|n|>\nu+1} O\big(\grave{P}^{01111}_{\nu mmnn}\big) z_m\bar z_m q_n\bar q_n \\
&\quad - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|\leqslant\nu} I_n^2 - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|=\nu+1}|z_n|^4 - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|>\nu+1}|q_n|^4 \\
&\quad + O\left(\varepsilon_0^6\left(\prod_{i=1}^{\nu-1}\varepsilon_i^2\right)\varepsilon_\nu^{38}\right) \quad (\text{lower order terms in } I, z, q) \\
&\quad + O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\big(|I|^2+|z|^5+|I||z|^3+|z|^3|q|^2+|I||q|^2+|q|^4\big) \\
&= \big\langle \omega^\nu\big(\xi(0),\dots,\xi(\nu)\big), I(\nu)\big\rangle + \sum_{|n|=\nu+1} \Omega_n^\nu\big(\xi(0),\dots,\xi(\nu)\big) z_n\bar z_n \\
&\quad + \sum_{|n|>\nu+1} \Omega_n^\nu\big(\xi(0),\dots,\xi(\nu)\big) q_n\bar q_n \\
&\quad + \sum_{|l|=1} O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right) I(\nu)^l + \sum_{|n|=\nu+1} O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right) z_n\bar z_n \\
&\quad + \sum_{|m|=|n|=\nu+1} O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right) z_m\bar z_m z_n\bar z_n \\
&\quad + \sum_{|l|=1,\,|n|=\nu+1} O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right) z_n\bar z_n I(\nu)^l + \sum_{|n|>\nu+1} O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right) q_n\bar q_n \\
&\quad + \sum_{|m|=\nu+1,\,|n|>\nu+1} O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right) z_m\bar z_m q_n\bar q_n
\end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|\leqslant\nu}I_n^2-\frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|=\nu+1}|z_n|^4-\frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|>\nu+1}|q_n|^4\\
&+O\left(\varepsilon_0^6\left(\prod_{i=1}^{\nu-1}\varepsilon_i^2\right)\varepsilon_\nu^{38}\right)\quad\text{(lower order terms in } I, z, q)\\
&+O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\left(|I|^2+|z|^5+|I||z|^3+|z|^3|q|^2+|I||q|^2+|q|^4\right).
\end{aligned}
$$

**Remark.** In the case of Dirichlet boundary conditions, $H_{\nu+1}$ should include one more term $O(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2)|z||q|^3$.

### 4.3. The new Hamiltonian

Below, we show that the new Hamiltonian $H_{\nu+1}$ enjoys similar properties as $H_\nu$. Let

$$
z_n=\sqrt{I_n+\xi_n}e^{\mathrm{i}\theta_n},\qquad \bar z_n=\sqrt{I_n+\xi_n}e^{-\mathrm{i}\theta_n},\quad |n|=\nu+1,
$$

and set $\xi(\nu+1)=(\ldots,\xi_n,\ldots)_{|n|=\nu+1}$, $I(\nu+1)=(\ldots,I_n,\ldots)_{|n|\leqslant\nu+1}$, $\theta(\nu+1)=(\ldots,\theta_n,\ldots)_{|n|\leqslant\nu+1}$, then

$$
\begin{aligned}
H_{\nu+1}={}&\left\langle\omega^\nu\big(\xi(0),\ldots,\xi(\nu)\big),I(\nu)\right\rangle+\sum_{|n|=\nu+1}\Omega_n^\nu\big(\xi(0),\ldots,\xi(\nu)\big)I_n\\
&+\sum_{|n|>\nu+1}\Omega_n^\nu\big(\xi(0),\ldots,\xi(\nu)\big)q_n\bar q_n\\
&+\sum_{|l|=1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I(\nu)^l+\sum_{|n|=\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I_n\\
&+\sum_{|m|=|n|=\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\xi_mI_n+\sum_{|m|=|n|=\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I_mI_n\\
&+\sum_{|l|=1,\,|n|=\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\xi_nI(\nu)^l+\sum_{|l|=1,\,|n|=\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I_nI(\nu)^l\\
&+\sum_{|n|>\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)q_n\bar q_n+\sum_{|m|=\nu+1,\,|n|>\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\xi_mq_n\bar q_n\\
&+\sum_{|m|=\nu+1,\,|n|>\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I_mq_n\bar q_n\\
&-\frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|\leqslant\nu}I_n^2-\frac{1}{2\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|=\nu+1}\xi_nI_n\\
&-\frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|=\nu+1}I_n^2-\frac{1}{4\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|>\nu+1}|q_n|^4
\end{aligned}
$$

$$+\,O\left(\varepsilon_0^6\left(\prod_{i=1}^{\nu-1}\varepsilon_i^2\right)\varepsilon_\nu^{38}\right)\quad\bigl(\text{lower order terms in } I(\nu+1),\xi(\nu+1),q\bigr)$$
$$+\,O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\Bigl(\bigl|I(\nu)\bigr|^2+\bigl|\xi(\nu+1)\bigr|^{\frac52}+\bigl|I(\nu)\bigr|\bigl|\xi(\nu+1)\bigr|^{\frac32}$$
$$+\bigl|\xi(\nu+1)\bigr|^{\frac32}|q|^2+\bigl|I(\nu)\bigr||q|^2+|q|^4\Bigr).$$

Let $\varepsilon_{\nu+1}=\varepsilon_\nu^{\frac95}$, consider the scalings: $\xi(\nu+1)\to\varepsilon_{\nu+1}^8\xi(\nu+1)$, $q\to\varepsilon_{\nu+1}^5 q$, and $I(\nu+1)\to\varepsilon_{\nu+1}^{10}I(\nu+1)$, we obtain the re-scaled Hamiltonian

$$\begin{aligned}
H_{\nu+1}&=\varepsilon_{\nu+1}^{-18}H_{\nu+1}\bigl(\varepsilon_{\nu+1}^8\xi(\nu+1),\varepsilon_{\nu+1}^{10}I(\nu+1),\varepsilon_{\nu+1}^5 q\bigr)\\
&=\bigl\langle\varepsilon_{\nu+1}^{-8}\omega^\nu\bigl(\xi(0),\dots,\xi(\nu)\bigr),I(\nu)\bigr\rangle+\sum_{|n|=\nu+1}\varepsilon_{\nu+1}^{-8}\Omega_n^\nu\bigl(\xi(0),\dots,\xi(\nu)\bigr)I_n\\
&\quad+\sum_{|n|>\nu+1}\varepsilon_{\nu+1}^{-8}\Omega_n^\nu\bigl(\xi(0),\dots,\xi(\nu)\bigr)q_n\bar q_n\\
&\quad+\sum_{|l|=1}\varepsilon_{\nu+1}^{-8}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I(\nu)^l+\sum_{|n|=\nu+1}\varepsilon_{\nu+1}^{-8}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I_n\\
&\quad+\sum_{|m|=|n|=\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\xi_m I_n+\sum_{|m|=|n|=\nu+1}\varepsilon_{\nu+1}^2 O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I_m I_n\\
&\quad+\sum_{|l|=1,\,|n|=\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\xi_n I(\nu)^l+\sum_{|l|=1,\,|n|=\nu+1}\varepsilon_{\nu+1}^2 O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I_n I(\nu)^l\\
&\quad+\sum_{|n|>\nu+1}\varepsilon_{\nu+1}^{-8}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)q_n\bar q_n+\sum_{|m|=\nu+1,\,|n|>\nu+1}O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\xi_m q_n\bar q_n\\
&\quad+\sum_{|m|=\nu+1,\,|n|>\nu+1}\varepsilon_{\nu+1}^2 O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)I_m q_n\bar q_n\\
&\quad-\frac{1}{4\pi}\varepsilon_{\nu+1}^2\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|\leqslant\nu}I_n^2-\frac{1}{2\pi}\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|=\nu+1}\xi_n I_n\\
&\quad-\frac{1}{4\pi}\varepsilon_{\nu+1}^2\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|=\nu+1}I_n^2-\frac{1}{4\pi}\varepsilon_{\nu+1}^2\left(\prod_{i=0}^{\nu}\varepsilon_i^2\right)\sum_{|n|>\nu+1}|q_n|^4\\
&\quad+\varepsilon_{\nu+1}^{-18}O\left(\varepsilon_0^6\left(\prod_{i=1}^{\nu-1}\varepsilon_i^2\right)\varepsilon_\nu^{38}\right)\quad\bigl(\text{lower order terms in } I(\nu+1),\xi(\nu+1),q\bigr)\\
&\quad+O\left(\varepsilon_0^6\prod_{i=1}^{\nu}\varepsilon_i^2\right)\bigl[\varepsilon_{\nu+1}^2\bigl|I(\nu)\bigr|^2+\varepsilon_{\nu+1}^2\bigl|\xi(\nu+1)\bigr|^{\frac52}+\varepsilon_{\nu+1}^4\bigl|I(\nu)\bigr|\bigl|\xi(\nu+1)\bigr|^{\frac32}\\
&\quad+\varepsilon_{\nu+1}^4\bigl|\xi(\nu+1)\bigr|^{\frac32}|q|^2+\varepsilon_{\nu+1}^2\bigl|I(\nu)\bigr||q|^2+\varepsilon_{\nu+1}^2|q|^4\bigr]
\end{aligned}$$

$$\begin{aligned}
&\stackrel{\text{def}}{=} \langle \omega^{\nu+1}(\xi(0),\dots,\xi(\nu+1)), I(\nu+1)\rangle + \sum_{|n|>\nu+1} \Omega_n^{\nu+1}(\xi(0),\dots,\xi(\nu+1)) q_n\bar{q}_n \\
&\quad - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu+1}\varepsilon_i^2\right)\sum_{|n|\leqslant\nu+1} I_n^2 - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu+1}\varepsilon_i^2\right)\sum_{|n|>\nu+1}|q_n|^4 + O\left(\varepsilon_0^6\left(\prod_{i=1}^{\nu+1}\varepsilon_i^2\right)\right) \\
&= N_{\nu+1} + \acute{P}_{\nu+1} + \grave{P}_{\nu+1} \\
&= N_{\nu+1} + P_{\nu+1},
\end{aligned}$$

where

$$N_{\nu+1} = \langle \omega^{\nu+1}(\xi(0),\dots,\xi(\nu+1)), I(\nu+1)\rangle + \sum_{|n|>\nu+1} \Omega_n^{\nu+1}(\xi(0),\dots,\xi(\nu+1)) q_n\bar{q}_n,$$

$$\xi(i) = (\dots,\xi_n,\dots)_{|n|=i}, \quad 0\leqslant i\leqslant \nu+1,$$

$$\omega^{\nu+1}(\xi(0),\dots,\xi(\nu+1)) = (\dots,\omega_n(\xi(0),\dots,\xi(\nu+1)),\dots)_{|n|\leqslant\nu+1},$$

$$\begin{aligned}
\omega_n(\xi(0),\dots,\xi(\nu+1)) &= \varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}\left(n^2+\frac{1}{5}\right) + \left(-\frac{1}{2\pi}\right)\varepsilon_0^2\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_\nu^{-8}\varepsilon_{\nu+1}^{-8}\xi_n \\
&\quad + O(\varepsilon_0^6\varepsilon_1^{-8}\cdots\varepsilon_{\nu+1}^{-8}) f_0^{\nu+1}(\xi(0)) + O(\varepsilon_0^6\varepsilon_2^{-8}\cdots\varepsilon_{\nu+1}^{-8}) f_1^{\nu+1}(\xi(1)) \\
&\quad + \cdots + O(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_{\nu+1}^{-8}) f_i^{\nu+1}(\xi(i)) \\
&\quad + \cdots + O(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{\nu-1}^2\varepsilon_{\nu+1}^{-8}) f_\nu^{\nu+1}(\xi(\nu)) \\
&\quad + O(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_\nu^2) f_{\nu+1}^{\nu+1}(\xi(\nu+1)), \quad |n|=i,
\end{aligned}$$

$$\left\| f_i^{\nu+1}(\xi(i))\right\|_{\mathcal{O}_\nu} = \sup_{\mathcal{O}_\nu}\left(\left|f_i^{\nu+1}(\xi(i))\right| + \left|\frac{\partial f_i^{\nu+1}}{\partial\xi(i)}\right|\right) \leqslant 1, \quad 0\leqslant i\leqslant\nu+1,$$

$$\begin{aligned}
\Omega_n^{\nu+1}(\xi(0),\dots,\xi(\nu+1)) &= \varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}\left(n^2+\frac{1}{5}\right) \\
&\quad + O(\varepsilon_0^6\varepsilon_1^{-8}\cdots\varepsilon_{\nu+1}^{-8}) f_0^{\nu+1}(\xi(0)) + O(\varepsilon_0^6\varepsilon_2^{-8}\cdots\varepsilon_{\nu+1}^{-8}) f_1^{\nu+1}(\xi(1)) \\
&\quad + \cdots + O(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_{\nu+1}^{-8}) f_i^{\nu+1}(\xi(i)) \\
&\quad + \cdots + O(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{\nu-1}^2\varepsilon_{\nu+1}^{-8}) f_\nu^{\nu+1}(\xi(\nu)) \\
&\quad + O(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_\nu^2) f_{\nu+1}^{\nu+1}(\xi(\nu+1)), \quad |n|>\nu+1,
\end{aligned}$$

$$P_{\nu+1} = \acute{P}_{\nu+1} + \grave{P}_{\nu+1},$$

$$\acute{P}_{\nu+1} = -\frac{1}{4\pi}\left(\prod_{i=0}^{\nu+1}\varepsilon_i^2\right)\sum_{|n|\leqslant\nu+1} I_n^2 - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu+1}\varepsilon_i^2\right)\sum_{|n|>\nu+1}|q_n|^4,$$

$$\grave{P}_{\nu+1} = O\left(\varepsilon_0^6\left(\prod_{i=1}^{\nu+1}\varepsilon_i^2\right)\right).$$

In Section 5.3, we will prove that, under the condition

$$\sum_{|n|\leqslant \nu+1} k_n n + \sum_{|n|=\nu+2} j_n n + \sum_{|n|>\nu+2} l_n n = 0,$$

there exists a subset $\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu \times [1,2]^2$ with $\operatorname{meas}(\mathcal{O}_\nu \times [1,2]^2 \setminus \mathcal{O}_{\nu+1}) < \varepsilon_{\nu+1}^{\frac{1}{20}}$, where

$$\mathcal{O}_{\nu+1} = \left\{ \big(\xi(0),\dots,\xi(\nu+1)\big)\colon \left|\langle k, \omega^{\nu+1}\rangle + \langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \geqslant \frac{\gamma_{\nu+1}}{|k|^{\tau_{\nu+1}}},\ \begin{array}{l} |k|+|j|+|l| \neq 0, \\ |j|+|l| \leqslant 4,\ |l| \leqslant 3 \end{array} \right\}$$

for $\gamma_{\nu+1} = (\prod_{i=0}^{\nu} \varepsilon_i)\varepsilon_{\nu+1}^{\frac{3}{4}}$, $\tau_{\nu+1} = 25(\nu+2)$, $\varepsilon_{\nu+1} = \varepsilon_\nu^{\frac{9}{5}}$ and $\bar{\Omega}^{\nu+1} = (\dots, \Omega_n^{\nu+1}, \dots)_{|n|=\nu+2}$, $\tilde{\Omega}^{\nu+1} = (\dots, \Omega_n^{\nu+1}, \dots)_{|n|>\nu+2}$.

**Remark.** In the case of Dirichlet boundary conditions, we will prove that,

$$|k| \leqslant K_{\nu+1}, \qquad \left| \sum_{n=\nu+2} \pm j_n n + \sum_{n>\nu+2} \pm l_n n \right| \leqslant K_{\nu+1},$$

there exists a subset $\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu \times [1,2]$ with $\operatorname{meas}(\mathcal{O}_\nu \times [1,2] \setminus \mathcal{O}_{\nu+1}) < \varepsilon_{\nu+1}^{\frac{1}{20}}$, where

$$\mathcal{O}_{\nu+1} = \left\{ \big(\xi(0),\dots,\xi(\nu+1)\big)\colon \left|\langle k, \omega^{\nu+1}\rangle + \langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \geqslant \frac{\gamma_{\nu+1}}{K_{\nu+1}^{\tau_{\nu+1}}},\ \begin{array}{l} |k|+|j|+|l| \neq 0, \\ |j|+|l| \leqslant 4,\ |l| \leqslant 3, \\ |j| = 0,\ \text{if } |l| = 3 \end{array} \right\}.$$

Moreover,

$$\|X_{\acute{P}_{\nu+1}}\|_{D_{\nu+1},\mathcal{O}_{\nu+1}} \leqslant \prod_{i=0}^{\nu+1} \varepsilon_i^2, \qquad \|X_{\grave{P}_{\nu+1}}\|_{D_{\nu+1},\mathcal{O}_{\nu+1}} \leqslant \varepsilon_0^6 \prod_{i=1}^{\nu+1} \varepsilon_i^2,$$

and thanks to Lemma 2.3, $P_{\nu+1} = \sum_{k,\alpha,\beta} P_{(\nu+1)k\alpha\beta}(I(\nu+1)) e^{\mathrm{i}\langle k, \theta(\nu+1)\rangle} q^\alpha \bar{q}^\beta$ has a compact form, i.e.,

$$P_{(\nu+1)k\alpha\beta} = 0, \quad \text{whenever} \sum_{|n|\leqslant \nu+1} k_n n + \sum_{|n|>\nu+1} (\alpha_n - \beta_n) n \neq 0.$$

This completes one step of KAM iterations.

## 5. Iteration lemma, convergence and measure estimates

For any given $s, r_0, \varepsilon_0$, $\xi(0) = (\dots, \xi_n, \dots)_{|n|=0}$, we define, for all $\nu \geqslant 0$, the following sequences

$$r_{\nu+1} = r_0 \left(1 - \sum_{i=2}^{\nu+2} 2^{-i}\right),$$

$$\varepsilon_{\nu+1} = \varepsilon_\nu^{\frac{9}{5}},$$

$$\gamma_\nu = \left(\prod_{i=0}^{\nu-1} \varepsilon_i\right)\varepsilon_\nu^{\frac{3}{4}}, \qquad \tau_\nu = 25(\nu+1),$$

$$\xi(\nu+1) = (\ldots, \xi_n, \ldots)_{|n|=\nu+1},$$

$$I(\nu) = (\ldots, I_n, \ldots)_{|n|\leqslant\nu},$$

$$\theta(\nu) = (\ldots, \theta_n, \ldots)_{|n|\leqslant\nu},$$

$$\omega^\nu\big(\xi(0), \ldots, \xi(\nu)\big) = \big(\ldots, \omega_n\big(\xi(0), \ldots, \xi(\nu)\big), \ldots\big)_{|n|\leqslant\nu},$$

$$\bar{\Omega}^\nu = \big(\ldots, \Omega_n^\nu, \ldots\big)_{|n|=\nu+1},$$

$$\tilde{\Omega}^\nu = \big(\ldots, \Omega_n^\nu, \ldots\big)_{|n|>\nu+1},$$

$$D_\nu = D(r_\nu, s),$$

$$\mathcal{O}_\nu = \left\{\big(\xi(0), \ldots, \xi(\nu)\big)\colon \left|\langle k, \omega^\nu\rangle + \langle j, \bar{\Omega}^\nu\rangle + \langle l, \tilde{\Omega}^\nu\rangle\right| \geqslant \frac{\gamma_\nu}{|k|^{\tau_\nu}}, \begin{array}{l} |k|+|j|+|l| \neq 0, \\ |j|+|l| \leqslant 4,\ |l| \leqslant 3 \end{array}\right\},$$

or $K_\nu = 2^\nu K_0$, $K_0 = |\ln\varepsilon_0|$ for the case of Dirichlet boundary conditions,

$$\mathcal{O}_\nu = \left\{\big(\xi(0), \ldots, \xi(\nu)\big)\colon \left|\langle k, \omega^\nu\rangle + \langle j, \bar{\Omega}^\nu\rangle + \langle l, \tilde{\Omega}^\nu\rangle\right| \geqslant \frac{\gamma_\nu}{K_\nu^{\tau_\nu}}, \begin{array}{l} |k|+|j|+|l| \neq 0, \\ |j|+|l| \leqslant 4,\ |l| \leqslant 3, \\ |j| = 0,\ \text{if } |l| = 3 \end{array}\right\}.$$

### *5.1. Iteration lemma*

The preceding analysis may be summarized as follows.

**Lemma 5.1.** *The following holds if $\varepsilon_0$ is sufficiently small. Suppose for any $\nu \geqslant 0$, $H_\nu = N_\nu + P_\nu = N_\nu + \acute{P}_\nu + \grave{P}_\nu$ is given on $D_\nu \times \mathcal{O}_\nu$ which real analytic in $(\theta(\nu), I(\nu), q, \bar{q}) \in D_\nu$ and Whitney smooth in $(\xi(0), \ldots, \xi(\nu)) \in \mathcal{O}_\nu$, where*

$$N_\nu = \big\langle\omega^\nu\big(\xi(0), \ldots, \xi(\nu)\big), I(\nu)\big\rangle + \sum_{|n|>\nu} \Omega_n^\nu\big(\xi(0), \ldots, \xi(\nu)\big) q_n \bar{q}_n,$$

*$P_\nu$ has a compact form, and*

$$\acute{P}_\nu = -\frac{1}{4\pi}\left(\prod_{i=0}^{\nu} \varepsilon_i^2\right)\sum_{|n|\leqslant\nu} I_n^2 - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu} \varepsilon_i^2\right)\sum_{|n|>\nu} |q_n|^4,$$

$$\grave{P}_\nu = O\left(\varepsilon_0^6\left(\prod_{i=1}^{\nu} \varepsilon_i^2\right)\right),$$

$$\|X_{\acute{P}_\nu}\|_{D_\nu, \mathcal{O}_\nu} \leqslant \prod_{i=0}^{\nu} \varepsilon_i^2, \qquad \|X_{\grave{P}_\nu}\|_{D_\nu, \mathcal{O}_\nu} \leqslant \varepsilon_0^6 \prod_{i=1}^{\nu} \varepsilon_i^2.$$

*Then there is a symplectic transformation*

$$\Phi_\nu : D_{\nu+1} \times \mathcal{O}_\nu \to D_\nu,$$

*which is real analytic in $(\theta(\nu), I(\nu), q, \bar{q}) \in D_{\nu+1}$ and Whitney smooth in $(\xi(0), \dots, \xi(\nu)) \in \mathcal{O}_\nu$, such that after the further action-angle transformation in the smaller domain, $H_{\nu+1} = N_{\nu+1} + P_{\nu+1} = N_{\nu+1} + \acute{P}_{\nu+1} + \grave{P}_{\nu+1}$ is defined on $D_{\nu+1} \times \mathcal{O}_{\nu+1}$ and enjoys similar properties as $H_\nu$, i.e., $N_{\nu+1}$ has the form*

$$N_{\nu+1} = \big\langle \omega^{\nu+1}\big(\xi(0), \dots, \xi(\nu+1)\big), I(\nu+1)\big\rangle + \sum_{|n|>\nu+1} \Omega_n^{\nu+1}\big(\xi(0), \dots, \xi(\nu+1)\big) q_n \bar{q}_n,$$

*$P_{\nu+1}$ has a compact form, and*

$$\acute{P}_{\nu+1} = -\frac{1}{4\pi}\left(\prod_{i=0}^{\nu+1} \varepsilon_i^2\right) \sum_{|n|\leqslant \nu+1} I_n^2 - \frac{1}{4\pi}\left(\prod_{i=0}^{\nu+1} \varepsilon_i^2\right) \sum_{|n|>\nu+1} |q_n|^4,$$

$$\grave{P}_{\nu+1} = O\left(\varepsilon_0^6\left(\prod_{i=1}^{\nu+1} \varepsilon_i^2\right)\right),$$

$$\|X_{\acute{P}_{\nu+1}}\|_{D_{\nu+1},\mathcal{O}_{\nu+1}} \leqslant \prod_{i=0}^{\nu+1} \varepsilon_i^2, \qquad \|X_{\grave{P}_{\nu+1}}\|_{D_{\nu+1},\mathcal{O}_{\nu+1}} \leqslant \varepsilon_0^6 \prod_{i=1}^{\nu+1} \varepsilon_i^2.$$

### 5.2. Convergence

Let $\Psi^\nu = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{\nu-1}$, $\nu = 1, 2, \dots$. Inductively, we have that $\Psi^\nu : D_{\nu+1} \times \mathcal{O}_\nu \to D_0$ and

$$H_0 \circ \Psi^\nu = H_\nu = N_\nu + P_\nu$$

for all $\nu \geqslant 1$.

Let $\tilde{\mathcal{O}} = \lim_{\nu\to\infty} \mathcal{O}_\nu$. We apply Lemma 5.1 and standard arguments (e.g. [26]) to conclude that $H_\nu, N_\nu, P_\nu, \Psi^\nu, D\Psi^\nu, \omega^\nu$ converge uniformly on $D(\frac{1}{2}r_0, 0) \times \tilde{\mathcal{O}}$, say to, $H_\infty, N_\infty, P_\infty, \Psi^\infty, D\Psi^\infty, \omega^\infty$ respectively. It is clear that

$$N_\infty = \big\langle \omega^\infty, I(\infty)\big\rangle.$$

Since

$$\varepsilon_\nu = \varepsilon_{\nu-1}^{\frac{9}{5}} = (\varepsilon_0)^{(\frac{9}{5})^\nu},$$

then if $\varepsilon_0$ is sufficiently small, we have by Lemma 5.1 that

$$X_{P_\infty}|_{D(\frac{1}{2}r_0,0)\times\tilde{\mathcal{O}}} \equiv 0.$$

Let $\phi_H^t$ denote the flow of any Hamiltonian vector field $X_H$. Since $H_0 \circ \Psi^\nu = H_\nu$, we have that

$$\phi_{H_0}^t \circ \Psi^\nu = \Psi^\nu \circ \phi_{H_\nu}^t.$$

The uniform convergence of $\Psi^\nu, D\Psi^\nu, X_{H_\nu}$ implies that one can pass the limit in the above to conclude that

$$\phi_{H_0}^t \circ \Psi^\infty = \Psi^\infty \circ \phi_{H_\infty}^t = \Psi^\infty \circ \phi_{N_\infty}^t$$

on $D(\frac{1}{2}r_0, 0) \times \tilde{\mathcal{O}}$. It follows that

$$\begin{aligned}&\phi^t_{H_0}\big(\Psi^\infty\big(\mathbb{T}^\infty\times\big\{\big(\xi(0),\xi(1),\ldots\big)\big\}\big)\big)\\&\quad=\Psi^\infty\phi^t_{N_\infty}\big(\mathbb{T}^\infty\times\big\{\big(\xi(0),\xi(1),\ldots\big)\big\}\big)=\Psi^\infty\big(\mathbb{T}^\infty\times\big\{\big(\xi(0),\xi(1),\ldots\big)\big\}\big)\end{aligned}$$

for all $(\xi(0),\xi(1),\ldots)\in\tilde{\mathcal{O}}$. Hence $\Psi^\infty(\mathbb{T}^\infty\times\{(\xi(0),\xi(1),\ldots)\})$ is an embedded invariant torus of the original perturbed Hamiltonian system at $(\xi(0),\xi(1),\ldots)\in\tilde{\mathcal{O}}$.

### 5.3. Measure estimate

According to Lemma 5.1, at the $\nu$th KAM step, we need to excise the parameter set $\mathcal{R}_{\nu+1}$ under the condition

$$\sum_{|n|\leqslant\nu+1}k_n n+\sum_{|n|=\nu+2}j_n n+\sum_{|n|>\nu+2}l_n n=0,$$

where

$$\begin{aligned}&\mathcal{R}_{\nu+1}\\&\quad=\left\{\big(\xi(0),\ldots,\xi(\nu+1)\big)\colon\ \big|\big\langle k,\omega^{\nu+1}\big\rangle+\big\langle j,\bar{\Omega}^{\nu+1}\big\rangle+\big\langle l,\tilde{\Omega}^{\nu+1}\big\rangle\big|<\frac{\gamma_{\nu+1}}{|k|^{\tau_{\nu+1}}},\ \begin{array}{l}|k|+|j|+|l|\neq0,\\|j|+|l|\leqslant4,\ |l|\leqslant3\end{array}\right\}.\end{aligned}$$

**Remark.** In the case of Dirichlet boundary conditions, at the $\nu$th KAM step, we need to excise the parameter set $\mathcal{R}_{\nu+1}$ under the condition

$$|k|\leqslant K_{\nu+1},\qquad\left|\sum_{n=\nu+2}\pm j_n n+\sum_{n>\nu+2}\pm l_n n\right|\leqslant K_{\nu+1},$$

where

$$\begin{aligned}&\mathcal{R}_{\nu+1}\\&\quad=\left\{\big(\xi(0),\ldots,\xi(\nu+1)\big)\colon\ \big|\big\langle k,\omega^{\nu+1}\big\rangle+\big\langle j,\bar{\Omega}^{\nu+1}\big\rangle+\big\langle l,\tilde{\Omega}^{\nu+1}\big\rangle\big|<\frac{\gamma_{\nu+1}}{K_{\nu+1}^{\tau_{\nu+1}}},\ \begin{array}{l}|k|+|j|+|l|\neq0,\\|j|+|l|\leqslant4,\ |l|\leqslant3\\|j|=0,\ \text{if }|l|=3\end{array}\right\}.\end{aligned}$$

**Lemma 5.2.** *When $k=0$, $\mathcal{R}_{\nu+1}=\emptyset$.*

**Proof.** When $k=0$, $0<|j|+|l|\leqslant4$,

$$\sum_{|n|=\nu+2}j_n n+\sum_{|n|>\nu+2}l_n n=0,$$

we consider $|\langle j,\bar{\Omega}^{\nu+1}\rangle+\langle l,\tilde{\Omega}^{\nu+1}\rangle|$. According to the number of minus signs in $(j,l)$, we distinguish them into the following five cases:

**Case 1.** *No minus sign.*

$$\big|\big\langle j,\bar{\Omega}^{\nu+1}\big\rangle+\big\langle l,\tilde{\Omega}^{\nu+1}\big\rangle\big|\geqslant\frac{1}{5}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}-O\big(\varepsilon_0^6\varepsilon_1^{-8}\cdots\varepsilon_{\nu+1}^{-8}\big)\geqslant\frac{1}{6}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.$$

**Case 2.** *One minus sign.* Under the condition $\sum_{|n|=\nu+2} j_n n + \sum_{|n|>\nu+2} l_n n = 0$,

$$\left|\langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \geqslant \frac{1}{5}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - O\left(\varepsilon_0^6\varepsilon_1^{-8}\cdots\varepsilon_{\nu+1}^{-8}\right) \geqslant \frac{1}{6}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.$$

**Case 3.** *Two minus signs.* If $0 < |j| + |l| < 4$, it can be handled in the same way as Cases 1 and 2. If $|j| + |l| = 4$, under the condition $\sum_{|n|=\nu+2} j_n n + \sum_{|n|>\nu+2} l_n n = 0$,

$$\left|\langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \geqslant \varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - O\left(\varepsilon_0^6\varepsilon_1^{-8}\cdots\varepsilon_{\nu+1}^{-8}\right) \geqslant \frac{1}{2}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.$$

**Case 4.** *Three minus signs.* In this case, the number of plus signs is at most 1, hence, under the condition $\sum_{|n|=\nu+2} j_n n + \sum_{|n|>\nu+2} l_n n = 0$,

$$\left|\langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \geqslant \frac{1}{5}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - O\left(\varepsilon_0^6\varepsilon_1^{-8}\cdots\varepsilon_{\nu+1}^{-8}\right) \geqslant \frac{1}{6}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.$$

**Case 5.** *Four minus signs.*

$$\left|\langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \geqslant \frac{1}{5}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - O\left(\varepsilon_0^6\varepsilon_1^{-8}\cdots\varepsilon_{\nu+1}^{-8}\right) \geqslant \frac{1}{6}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.$$

In conclusion, Lemma 5.2 follows. □

**Remark.** In the case of Dirichlet boundary conditions, because we have the further restriction $|j| = 0$ for $|l| = 3$, hence Lemma 5.2 still holds true.

**Lemma 5.3.** *When $k \neq 0$, $|l| = 1$, $|\langle l, \tilde{\Omega}^{\nu+1}\rangle| = |\Omega_n^{\nu+1}| \geqslant \frac{1}{2}n^2\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}$, then when $|n| > 100(\nu+2)|k|$, $\mathcal{R}_{\nu+1} = \emptyset$.*

**Proof.**

$$\begin{aligned}
&\left|\langle k, \omega^{\nu+1}\rangle + \langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \\
&\quad \geqslant \frac{1}{2}n^2\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - 6\big((\nu+2)\big)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant \frac{1}{2}\big(100(\nu+2)|k|\big)^2\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - 6\big((\nu+2)\big)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant (\nu+2)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.
\end{aligned}$$

Hence, Lemma 5.3 is obtained. □

**Lemma 5.4.** *When $k \neq 0$, $|l| = 2$, $|\langle l, \tilde{\Omega}^{\nu+1}\rangle| = |\Omega_n^{\nu+1} \pm \Omega_m^{\nu+1}|$, then when $\max\{|n|, |m|\} > 100(\nu+2)^2|k|$, $\mathcal{R}_{\nu+1} = \emptyset$.*

**Proof.** When $m = -n$ and $|\langle l, \tilde{\Omega}^{\nu+1}\rangle| = |\Omega_n^{\nu+1} - \Omega_m^{\nu+1}|$, then due to $\max\{|n|, |m|\} > 100(\nu+2)^2|k|$,

$$\left|\sum_{|n|\leqslant\nu+1} k_n n + \sum_{|n|=\nu+2} j_n n + \sum_{|n|>\nu+2} l_n n\right| \geqslant 100(\nu+2)^2|k| \neq 0,$$

i.e., $\mathcal{R}_{\nu+1} = \emptyset$. When $|m| \neq |n|$,

$$\begin{aligned}
&\left|\langle k, \omega^{\nu+1}\rangle + \langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \\
&\quad \geqslant \frac{1}{2}\max\{|n|, |m|\}\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - 6\big((\nu+2)\big)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant \frac{1}{2}\big(100(\nu+2)^2|k|\big)\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - 6\big((\nu+2)\big)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant (\nu+2)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.
\end{aligned}$$

Hence, Lemma 5.4 is also obtained. □

**Lemma 5.5.** *When* $k \neq 0$, $|l| = 3$, $|\langle l, \tilde{\Omega}^{\nu+1}\rangle| = |\Omega_p^{\nu+1} \pm \Omega_n^{\nu+1} \pm \Omega_m^{\nu+1}|$, *then when* $\max\{|p|, |n|, |m|\} > 4(100(\nu+2)^2|k|)^2$, $\mathcal{R}_{\nu+1} = \emptyset$.

**Proof.** Without loss of generality, we may assume $|p| \leqslant |n| \leqslant |m|$. Then when there are the same signs in the front of $\Omega_n^{\nu+1}$ and $\Omega_m^{\nu+1}$, we have

$$\begin{aligned}
&\left|\langle k, \omega^{\nu+1}\rangle + \langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \\
&\quad = \left|\langle k, \omega^{\nu+1}\rangle + \langle j, \bar{\Omega}^{\nu+1}\rangle \pm \Omega_p^{\nu+1} \pm \left(\Omega_n^{\nu+1} + \Omega_m^{\nu+1}\right)\right| \\
&\quad \geqslant \frac{1}{2}\left(m^2 + \frac{1}{5}\right)\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - 6\big((\nu+2)\big)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant \big(100(\nu+2)^2|k|\big)^2\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - 6\big((\nu+2)\big)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant (\nu+2)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.
\end{aligned}$$

If there are the different signs in the front of $\Omega_n^{\nu+1}$ and $\Omega_m^{\nu+1}$, we distinguish it into the following three cases:

**Case 1.** $|p| \leqslant 100(\nu+2)^2|k|$ and $m = -n$. At this time,

$$\begin{aligned}
\left|\sum_{|n|\leqslant\nu+1} k_n n + \sum_{|n|=\nu+2} j_n n \pm p \pm (n-m)\right| &\geqslant |n| - |p| - 6(\nu+2)|k| \\
&\geqslant \big(100(\nu+2)^2|k|\big)^2 - 100(\nu+2)^2|k| - 6(\nu+2)|k| \\
&\neq 0,
\end{aligned}$$

i.e., $\mathcal{R}_{\nu+1} = \emptyset$.

**Case 2.** $|p| \leqslant 100(\nu+2)^2|k|$ and $|m| \neq |n|$. Then due to $|m| > 4(100(\nu+2)^2|k|)^2$, one can get

$$\begin{aligned}
&\left|\langle k, \omega^{\nu+1}\rangle + \langle j, \bar{\Omega}^{\nu+1}\rangle + \langle l, \tilde{\Omega}^{\nu+1}\rangle\right| \\
&\quad \geqslant \frac{1}{2}|m|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - |p|^2\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - 6(\nu+2)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant 2\big(100(\nu+2)^2|k|\big)^2\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - \big(100(\nu+2)^2|k|\big)^2\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - 6(\nu+2)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant (\nu+2)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.
\end{aligned}$$

**Case 3.** $|p| > 100(\nu+2)^2|k|$. Then $m = n \pm p + \sum_{|n|\leqslant\nu+1} k_n n + \sum_{|n|=\nu+2} j_n n$, substitute it into $-\Omega_m^{\nu+1} + \Omega_n^{\nu+1} \pm \Omega_p^{\nu+1} + \langle k, \omega_{\nu+1}\rangle + \langle j, \bar{\Omega}^{\nu+1}\rangle$, then one has

$$
\begin{aligned}
&\bigl|-\Omega_m^{\nu+1} + \Omega_n^{\nu+1} \pm \Omega_p^{\nu+1} + \bigl\langle k, \omega^{\nu+1}\bigr\rangle + \bigl\langle j, \bar{\Omega}^{\nu+1}\bigr\rangle\bigr| \\
&\quad \geqslant |p||n|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} - 6(\nu+2)^2|k|\bigl(|p|+|n|\bigr)\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant \bigl(|p| - 12(\nu+2)^2|k|\bigr)|n|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant \bigl(100(\nu+2)^2|k| - 12(\nu+2)^2|k|\bigr)100(\nu+2)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8} \\
&\quad \geqslant (\nu+2)^2|k|\varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}.
\end{aligned}
$$

Hence, Lemma 5.5 is proved. □

**Remark.** In the case of Dirichlet boundary conditions, without loss of generality, we may assume $p < n < m$, otherwise, they can be combined into the cases $|l| = 1$ or $|l| = 2$. The most complicated case is $p > 100(\nu+2)^2 K_{\nu+1}$, according to the truncation $|m-n-p| \leqslant K_{\nu+1}$, then $m > n + p - K_{\nu+1}$, similar to the above proof, we can get Lemma 5.5.

**Lemma 5.6.** *For a fixed $k \neq 0$, $j$, $l$, when the frequencies, as a function of $\xi$, can be regarded as the independent parameters, one has*

$$
\operatorname{meas}\bigl(\mathcal{R}_{\nu+1}^{k,j,l}\bigr) \leqslant \frac{\varepsilon_{\nu+1}^{\frac{1}{18}}}{|k|^{\tau_{\nu+1}}}.
$$

**Proof.** Let us recall

$$
\xi(i) = (\ldots, \xi_n, \ldots)_{|n|=i}, \quad 0 \leqslant i \leqslant \nu+1,
$$

$$
\omega^{\nu+1}\bigl(\xi(0), \ldots, \xi(\nu+1)\bigr) = \bigl(\ldots, \omega_n\bigl(\xi(0), \ldots, \xi(\nu+1)\bigr), \ldots\bigr)_{|n|\leqslant\nu+1},
$$

$$
\begin{aligned}
\omega_n\bigl(\xi(0), \ldots, \xi(\nu+1)\bigr) = {}& \varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}\left(n^2 + \frac{1}{5}\right) + \left(-\frac{1}{2\pi}\right)\varepsilon_0^2\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_\nu^{-8}\varepsilon_{\nu+1}^{-8}\xi_n \\
&+ O\bigl(\varepsilon_0^6\varepsilon_1^{-8}\cdots\varepsilon_{\nu+1}^{-8}\bigr) f_0^{\nu+1}\bigl(\xi(0)\bigr) + O\bigl(\varepsilon_0^6\varepsilon_2^{-8}\cdots\varepsilon_{\nu+1}^{-8}\bigr) f_1^{\nu+1}\bigl(\xi(1)\bigr) \\
&+ \cdots + O\bigl(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_{\nu+1}^{-8}\bigr) f_i^{\nu+1}\bigl(\xi(i)\bigr) \\
&+ \cdots + O\bigl(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{\nu-1}^2\varepsilon_{\nu+1}^{-8}\bigr) f_\nu^{\nu+1}\bigl(\xi(\nu)\bigr) \\
&+ O\bigl(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_\nu^2\bigr) f_{\nu+1}^{\nu+1}\bigl(\xi(\nu+1)\bigr), \quad |n| = i,
\end{aligned}
$$

$$
\bigl\|f_i^{\nu+1}\bigl(\xi(i)\bigr)\bigr\|_{\mathcal{O}_\nu} = \sup_{\mathcal{O}_\nu}\left(\bigl|f_i^{\nu+1}\bigl(\xi(i)\bigr)\bigr| + \left|\frac{\partial f_i^{\nu+1}}{\partial\xi(i)}\right|\right) \leqslant 1, \quad 0 \leqslant i \leqslant \nu+1,
$$

$$
\begin{aligned}
\Omega_n^{\nu+1}\bigl(\xi(0), \ldots, \xi(\nu+1)\bigr) = {}& \varepsilon_0^{-8}\cdots\varepsilon_{\nu+1}^{-8}\left(n^2 + \frac{1}{5}\right) \\
&+ O\bigl(\varepsilon_0^6\varepsilon_1^{-8}\cdots\varepsilon_{\nu+1}^{-8}\bigr) f_0^{\nu+1}\bigl(\xi(0)\bigr) + O\bigl(\varepsilon_0^6\varepsilon_2^{-8}\cdots\varepsilon_{\nu+1}^{-8}\bigr) f_1^{\nu+1}\bigl(\xi(1)\bigr) \\
&+ \cdots + O\bigl(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_{\nu+1}^{-8}\bigr) f_i^{\nu+1}\bigl(\xi(i)\bigr)
\end{aligned}
$$

$$+\cdots+O\left(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{\nu-1}^2\varepsilon_{\nu+1}^{-8}\right)f_\nu^{\nu+1}\big(\xi(\nu)\big)$$
$$+O\left(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_\nu^2\right)f_{\nu+1}^{\nu+1}\big(\xi(\nu+1)\big),\quad |n|>\nu+1.$$

According to the assumption, the map $(\xi(0),\ldots,\xi(\nu+1))\to\omega^{\nu+1}(\xi(0),\ldots,\xi(\nu+1))$ is a diffeomorphism, and

$$\left|\frac{\partial\omega_n}{\partial\xi_n}\right|\geqslant\left(\frac{1}{2\pi}\right)\varepsilon_0^2\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_\nu^{-8}\varepsilon_{\nu+1}^{-8}-O\left(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_{\nu+1}^{-8}\right)$$
$$\geqslant\left(\frac{1}{4\pi}\right)\varepsilon_0^2\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_\nu^{-8}\varepsilon_{\nu+1}^{-8},\quad 0\leqslant|n|=i\leqslant\nu+1,$$

moreover, due to $k=(\ldots,k_n,\ldots)_{|n|\leqslant\nu+1}\neq 0$, then there exists the maximum valued component $|k_n|$. If $|n|=i$, $0\leqslant i\leqslant\nu+1$, then

$$\left|\frac{\partial(\langle k,\omega^{\nu+1}\rangle+\langle j,\bar{\Omega}^{\nu+1}\rangle+\langle l,\tilde{\Omega}^{\nu+1}\rangle)}{\partial\xi_n}\right|$$
$$\geqslant\left(\frac{1}{4\pi}\right)|k_n|\varepsilon_0^2\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_\nu^{-8}\varepsilon_{\nu+1}^{-8}-|k_n|O\left(\varepsilon_0^6\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_{\nu+1}^{-8}\right)$$
$$\geqslant\left(\frac{1}{8\pi}\right)|k_n|\varepsilon_0^2\varepsilon_1^2\cdots\varepsilon_{i-1}^2\varepsilon_{i+1}^{-8}\cdots\varepsilon_\nu^{-8}\varepsilon_{\nu+1}^{-8},$$

then

$$\operatorname{meas}\big(\mathcal{R}_{\nu+1}^{k,j,l}\big)\leqslant\frac{\gamma_{\nu+1}}{|k|^{\tau_{\nu+1}}}\frac{1}{\varepsilon_0^2\varepsilon_1^2\cdots\varepsilon_\nu^2}\leqslant\frac{\prod_{i=0}^{\nu}\varepsilon_i\varepsilon_{\nu+1}^{\frac{3}{4}}}{|k|^{\tau_{\nu+1}}}\frac{1}{\varepsilon_0^2\varepsilon_1^2\cdots\varepsilon_\nu^2}$$
$$\leqslant\frac{\varepsilon_{\nu+1}^{\frac{3}{4}}}{|k|^{\tau_{\nu+1}}}\frac{1}{\varepsilon_\nu^{\frac{5}{4}}}\leqslant\frac{\varepsilon_{\nu+1}^{\frac{3}{4}}}{|k|^{\tau_{\nu+1}}}\frac{1}{\varepsilon_{\nu+1}^{\frac{25}{36}}}\leqslant\frac{\varepsilon_{\nu+1}^{\frac{1}{18}}}{|k|^{\tau_{\nu+1}}}.\qquad\square$$

**Lemma 5.7.**

$$\operatorname{meas}\Big(\bigcup_{\nu\geqslant 0}\mathcal{R}_{\nu+1}\Big)=\operatorname{meas}\Big(\bigcup_{\nu\geqslant 0}\bigcup_{k,j,l}\mathcal{R}_{\nu+1}^{k,j,l}\Big)\leqslant\varepsilon_0^{\frac{1}{20}}.$$

**Proof.** Due to Lemma 5.6, we have $\operatorname{meas}(\mathcal{R}_{\nu+1}^{k,j,l})\leqslant\frac{\varepsilon_{\nu+1}^{\frac{1}{18}}}{|k|^{\tau_{\nu+1}}}$. By Lemmas 5.2–5.5, one can get that

$$\operatorname{meas}\Big(\bigcup_{\nu\geqslant 0}\mathcal{R}_{\nu+1}\Big)=\operatorname{meas}\Big(\bigcup_{\nu\geqslant 0}\bigcup_{k,j,l}\mathcal{R}_{\nu+1}^{k,j,l}\Big)$$
$$\leqslant\sum_{\nu\geqslant 0}\sum_{k}\big[4\big(100(\nu+2)^2|k|\big)^2\big]^4\frac{\varepsilon_{\nu+1}^{\frac{1}{18}}}{|k|^{\tau_{\nu+1}}}$$

$$\leqslant \sum_{\nu\geqslant 0}\sum_{k}|k|^8\frac{\varepsilon_{\nu+1}^{\frac{1}{20}}}{|k|^{\tau_{\nu+1}}}\leqslant \sum_{\nu\geqslant 0}\sum_{l\geqslant 1}l^{2(\nu+1)+1}\frac{\varepsilon_{\nu+1}^{\frac{1}{20}}}{l^{25(\nu+2)}}$$

$$\leqslant \sum_{\nu\geqslant 0}\varepsilon_{\nu+1}^{\frac{1}{20}}\leqslant \varepsilon_0^{\frac{1}{20}}.$$

This completes the measure estimate. □

**Remark.** In the case of Dirichlet boundary conditions, we need to change the above $|k|$ into $K_{\nu+1}$.

## Acknowledgments

Part of this work was carried out during the author's stay at Department of Mathematics and Statistics, McMaster University. The author is grateful to Walter Craig for reading this paper patiently and improving this paper's English presentation, the author is also grateful to Dario Bambusi for expressing his great interest to this paper. The author would like to thank the referees for their valuable comments and suggestions which help to improve this paper.

## References

[1] D. Bambusi, On long time stability in Hamiltonian perturbations of non-resonant linear PDEs, Nonlinearity 12 (1999) 823–850.
[2] J. Bourgain, Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, Int. Math. Res. Not. IMRN (1994) 475–497.
[3] J. Bourgain, Construction of periodic solutions of nonlinear wave equations in higher dimension, Geom. Funct. Anal. 5 (1995) 629–639.
[4] J. Bourgain, Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, Ann. of Math. 148 (1998) 363–439.
[5] J. Bourgain, Green's Function Estimates for Lattice Schrödinger Operators and Applications, Ann. of Math. Stud., vol. 158, Princeton University Press, Princeton, NJ, 2005.
[6] J. Bourgain, Nonlinear Schrödinger Equations, IAS/Park City Math. Ser., vol. 5, American Mathematical Society, Providence, RI, 1999.
[7] J. Bourgain, Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations, Geom. Funct. Anal. 6 (1996) 201–230.
[8] J. Bourgain, On invariant tori of full dimension for 1D periodic NLS, J. Funct. Anal. 229 (2005) 62–94.
[9] L. Chierchia, J. You, KAM tori for 1D nonlinear wave equations with periodic boundary conditions, Comm. Math. Phys. 211 (2000) 498–525.
[10] W. Craig, C.E. Wayne, Newton's method and periodic solutions of nonlinear wave equations, Comm. Pure Appl. Math. 46 (1993) 1409–1498.
[11] L.H. Eliasson, Perturbations of stable invariant tori for Hamiltonian systems, Ann. Sc. Norm. Super. Pisa Cl. Sci. 15 (1988) 115–147.
[12] L.H. Eliasson, S.B. Kuksin, KAM for non-linear Schrödinger equation, Ann. of Math. 172 (2010) 371–435.
[13] J. Geng, X. Xu, J. You, An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation, Adv. Math. 226 (2011) 5361–5402.
[14] J. Geng, J. You, A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions, J. Differential Equations 209 (2005) 1–56.
[15] J. Geng, J. You, A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces, Comm. Math. Phys. 262 (2006) 343–372.
[16] J. Geng, J. You, KAM tori for higher dimensional beam equations with constant potentials, Nonlinearity 19 (2006) 2405–2423.
[17] J. Geng, Y. Yi, Quasi-periodic solutions in a nonlinear Schrödinger equation, J. Differential Equations 233 (2007) 512–542.
[18] T. Kappeler, J. Pöschel, KdV & KAM, Springer, 2003.
[19] S.B. Kuksin, Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum, Funct. Anal. Appl. 21 (1987) 192–205.
[20] S.B. Kuksin, Nearly Integrable Infinite Dimensional Hamiltonian Systems, Lecture Notes in Math., vol. 1556, Springer, Berlin, 1993.
[21] S.B. Kuksin, A KAM-theorem for equations of the Korteweg–de Vries type, Rev. Math. Phys. 10 (1998) 1–64.
[22] S.B. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasiperiodic oscillations for a nonlinear Schrödinger equation, Ann. of Math. 143 (1996) 149–179.

[23] Z. Liang, J. You, Quasi-periodic solutions for 1D Schrödinger equations with higher order nonlinearity, SIAM J. Math. Anal. 36 (2005) 1965–1990.
[24] H. Niu, J. Geng, Almost periodic solutions for a class of higher dimensional beam equations, Nonlinearity 20 (2007) 2499–2517.
[25] J. Pöschel, On elliptic lower dimensional tori in Hamiltonian systems, Math. Z. 202 (1989) 559–608.
[26] J. Pöschel, Quasi-periodic solutions for a nonlinear wave equation, Comment. Math. Helv. 71 (1996) 269–296.
[27] J. Pöschel, A KAM theorem for some nonlinear partial differential equations, Ann. Sc. Norm. Super. Pisa Cl. Sci. 23 (1996) 119–148.
[28] J. Pöschel, On the construction of almost periodic solutions for a nonlinear Schrödinger equation, Ergodic Theory Dynam. Systems 22 (2002) 1537–1549.
[29] J. Pöschel, E. Trubowitz, Inverse Spectral Theory, Academic Press, New York, 1987.
[30] C.E. Wayne, Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory, Comm. Math. Phys. 127 (1990) 479–528.
[31] J. Xu, Q. Qiu, J. You, A KAM theorem of degenerate infinite dimensional Hamiltonian systems (I), Sci. China Ser. A 39 (1996) 372–383.
[32] J. Xu, Q. Qiu, J. You, A KAM theorem of degenerate infinite dimensional Hamiltonian systems (II), Sci. China Ser. A 39 (1996) 384–394.
[33] X. Yuan, Quasi-periodic solutions of completely resonant nonlinear wave equations, J. Differential Equations 230 (2006) 213–274.